# Deep Images of the Galactic Center with GRAVITY

GRAVITY Collaboration[⋆]: R. Abuter[8], N. Aimar[2], A. Amorim[6,12], P. Arras[17,26], M. Bauböck[1,18], J.P. Berger[5,8], H. Bonnet[8], W. Brandner[3], G. Bourdarot[5,1], V. Cardoso[12,20], Y. Clénet[2], R. Davies[1], P.T. de Zeeuw[10,1], J. Dexter[13,1], Y. Dallilar[1], A. Drescher[1], F. Eisenhauer[1], T. Enßlin[17], N.M. Förster Schreiber[1], P. Garcia[7,12], F. Gao[1,19], E. Gendron[2], R. Genzel[1,11], S. Gillessen[1], M. Habibi[1], X. Haubois[9], G. Heißel[2], T. Henning[3], S. Hippler[3], M. Horrobin[4], A. Jiménez-Rosales[1,21], L. Jochum[9], L. Jocou[5], A. Kaufer[9], P. Kervella[2], S. Lacour[2], V. Lapeyrère[2], J.-B. Le Bouquin[5], P. Léna[2], D. Lutz[1], F. Mang[1], M. Nowak[15,2], T. Ott[1], T. Paumard[2], K. Perraut[5], G. Perrin[2], O. Pfuhl[8,1], S. Rabien[1], J. Shangguan[1], T. Shimizu[1], S. Scheithauer[3], J. Stadler[1,17], O. Straub[1], C. Straubmeier[4], E. Sturm[1], L.J. Tacconi[1], K.R.W. Tristram[9], F. Vincent[2], S. von Fellenberg[1], I. Waisberg[14,1], F. Widmann[1], E. Wieprecht[1], E. Wiezorrek[1], J. Woillez[8], S. Yazici[1,4], A. Young[1], and G. Zins[9]

*(Affiliations can be found after the references)*

December 13, 2021

**ABSTRACT**

Stellar orbits at the Galactic Center provide a very clean probe of the gravitational potential of the supermassive black hole. They can be studied with unique precision, beyond the confusion limit of a single telescope, with the near-infrared interferometer GRAVITY. Imaging is essential to search the field for faint, unknown stars on short orbits which potentially could constrain the black hole spin. Furthermore, it provides the starting point for astrometric fitting to derive highly accurate stellar positions. Here, we present $G^R$, a new imaging tool specifically designed for Galactic Center observations with GRAVITY. The algorithm is based on a Bayesian interpretation of the imaging problem, formulated in the framework of information field theory and building upon existing works in radio-interferometric imaging. Its application to GRAVITY observations from 2021 yields the deepest images to date of the Galactic Center on scales of a few milliarcseconds. The images reveal the complicated source structure within the central 100 mas around Sgr A*, where we detected the stars S29 and S55 and confirm S62 on its trajectory, slowly approaching Sgr A*. Furthermore, we were able to detect S38, S42, S60, and S63 in a series of exposures for which we offset the fiber from Sgr A*. We provide an update on the orbits of all aforementioned stars. In addition to these known sources, the images also reveal a faint star moving to the west at a high angular velocity. We cannot find any coincidence with any known source and, thus, we refer to the new star as S300. From the flux ratio with S29, we estimate its K-band magnitude as $m_K$ (S300) $\simeq 19.0 - 19.3$. Images obtained with CLEAN confirm the detection. To assess the sensitivity of our images, we note that fiber damping reduces the apparent magnitude of S300 and the effect increases throughout the year as the star moves away from the field center. Furthermore, we performed a series of source injection tests. Under favorable circumstances, sources well below a magnitude of 20 can be recovered, while 19.7 is considered the more universal limit for a good data set.

**Key words.** Black hole physics, Galaxy: nucleus, Techniques: image processing, Techniques: high angular resolution, Methods: numerical, Methods: statistical

## 1. Introduction

The Galactic Center (GC) is a unique laboratory for probing general relativity (GR) (Genzel et al. 2010), where stars orbiting Sagittarius A* (Sgr A*) serve as clean test particles in the gravitational field of a supermassive black hole (SMBH). With near-infrared (near-IR) interferometry, it is not only possible to peer through the dust obscuring the GC, but to also push the angular resolution beyond a single telescope's diffraction limit. Down to an astrometric accuracy of ∼ 65 μas, this technique has been pioneered by the GRAVITY instrument (Gravity Collaboration et al. 2017), which couples the four 8m telescopes at the ESO Very Large Telescope (VLT). At the available telescope separation of ≲ 130 m, Sgr A* and the stars in its vicinity still appear

as point sources, but their positions can be determined by a factor of ∼ 20 more accurately compared to adaptive optics (AO) assisted imaging with a single telescope of similar size. Most importantly, the high angular resolution of GRAVITY allows to overcome the confusion limit of AO imaging.

There are about 50 stars with known orbits within 1 arc second (as) of Sgr A* (Gillessen et al. 2017). The most prominent of them, S2, passed its pericenter in 2018. The close monitoring of this event allowed for the detection of the gravitational redshift (Gravity Collaboration et al. 2018a; Do et al. 2019) and the Schwarzschild precession (Gravity Collaboration et al. 2020b) in the S2 orbit. In 2019, S2 still was the brightest source next to Sgr A* within the GRAVITY field of view (FOV) (Gravity Collaboration et al. 2021a); in 2021, it moved to a separation that is sufficiently large such that the two sources cannot be detected in a single pointing. Two other stars, however, approach Sgr A* and go through their pericenters in 2021 – S29 and S55 (Gillessen et al. 2017), which we have detected within 50 mas from Sgr A*. In comparison with S2, S29 approaches the black hole even more closely, while S55 has a shorter orbital pe-

riod (Meyer et al. 2012). We present updated results from their GR orbits in a second paper (Gravity Collaboration 2021).

The magnitude at which gravitational redshift and Schwarzschild precession affect stellar orbits are on the order of $\beta^2$ (where $\beta = v/c$, Zucker et al. 2006), while the Lense-Thirring precession due to the black hole spin falls off faster with the distance to the black hole. It is thus not clear whether any of the known stellar orbits allow for the detection of higher-order GR effects (Merritt et al. 2010; Zhang & Iorio 2017). A faint star at a smaller radius, on the other hand, could provide an opportunity to measure the spin of the SMBH (Waisberg et al. 2018). The expected number of stars suitable for such a measurement has been estimated around unity from extrapolation of the density profile and mass function observed at the GC (Genzel et al. 2003b; Do et al. 2013; Gallego-Cano et al. 2018) to small radii and faint stars, respectively (Waisberg et al. 2018).

With the interferometry approach, each baseline between two telescopes represents a point in Fourier space and the correlation of their signal corresponds to the Fourier transform of the image at this coordinate, in accordance with the well-known van Cittert-Zernike theorem (van Cittert 1934; Zernike 1938). Taking advantage of the Earth's rotation over a longer observing sequence and multi-wavelength observations help to fill the so-called $(u, v)$-plane. If sufficient prior knowledge on the observed flux distribution exists, model fitting is a powerful method to extract the desired information from the data. Without any prior indication of where a source might be expected, on the other hand, image reconstruction is the technique of choice in the search for faint, as-yet-unknown stars.

Consequently, deep imaging is essential in pushing the exploration of the GC further. Beyond the quest for faint stars, it is also the method of choice for exploring a lesser known field and serves as the starting point for astrometric fitting. The detection of S62, a slowly moving star at the K-band magnitude of $m_K$ (S62) $\simeq 18.9$, in GRAVITY images reconstructed with the radio-interferometry algorithm CLEAN (Högbom 1974), clearly demonstrates the power and value of the imaging approach (Gravity Collaboration et al. 2021a). CLEAN views the image as a collection of point sources, whose signal it subtracts iteratively from the measured coherent flux until only the noise is left. The question of where to place those point sources and when to stop iterations is guided by the Fourier inversion of the data. Because the $(u, v)$-space is only sparsely sampled, this "dirty image" is usually dominated by the inverse Fourier transform of the sampling pattern (the so-called dirty beam) and only the most prominent sources in the field are apparent. After subtracting their signal, fainter sources become recognizable in the residual images. As such, CLEAN depends on the linearity and invertibility of the Fourier transform that relates the image to the data.

Both conditions are not strictly satisfied for near-IR interferometry, where instrumental and observational effects, such as the finite bandwidth size and optical aberrations introduced by the instrument (Gravity Collaboration et al. 2021b), complicate the measurement equation. This is similar to what is known as direction-dependent effects (DDEs) in radio interferometric imaging (see e.g. Bhatnagar et al. 2008; Smirnov 2011). Furthermore, the requirement for a linear measurement equation impedes the use of closure quantities (Rogers et al. 1974), non-linear combinations of the data on different baselines that are insensitive to any telescope-based errors and thus provide more robust measurements.

Bayesian forward modeling offers an alternative approach to image reconstruction in which possible flux distributions are

quantified by the prior and the optimal image is found from exploration of the joint prior and likelihood, that is, the posterior. In this fashion, only the forward implementation of the measurement equation (and possibly its derivative for effective minimization) are needed, such that non-linear and non-invertible terms can be handled straightforwardly. The method, however, comes at the disadvantage of increased computational costs required to perform the posterior search.

Existing tools for optical and near-IR interferometric imaging, such as MIRA (Thiébaut 2008) and SQUEZZE (Baron et al. 2010), perform the posterior search either by descent minimization or with Monte Carlo Markov Chains (MCMC). While it may be faster, the former method is limited to convex likelihood and prior formulations. The MCMC method, on the other hand, becomes very inefficient for high-dimensional problems. For imaging, where every pixel is a free parameter to be inferred, this limits the applicability to rather small grids. Furthermore, these are general-purpose codes intended for the application to a range of instruments and, therefore, they do not account for all instrumental effects known to impact GRAVITY data. Finally, imaging the GC is complicated by the variability of Sgr A* on a timescale of five minutes (Genzel et al. 2003a; Ghez et al. 2004; Eisenhauer et al. 2005; Gillessen et al. 2006; Eckart et al. 2008; Do et al. 2009; Dodds-Eden et al. 2011; Witzel et al. 2018; Gravity Collaboration et al. 2020a,c). Naively combined over a longer period, the data become inconsistent, but $(u, v)$-sparsity prevents snapshot images over such a short duration. Rather, a prior model beyond those currently available in public codes is required, which can accommodate flux variations of the central source in a otherwise static field of point sources.

In this paper, we present our new imaging code GRAVITY-RESOLVE $(G^R)$[1]. It builds upon RESOLVE (Arras et al. 2021a; Arras et al. 2018), a Bayesian algorithm for radio interferometry, but it is tailored to GC observations with GRAVITY in its measurement equation and its prior model. For exploring the posterior distribution, we employed Metric Gaussian Variational Inference (MGVI, Knollmüller & Enßlin 2019), an algorithm that aims to provide a trade-off between robustness to complicated posterior shapes and applicability to high-dimensional problems. We present the details of $G^R$ in Sect. 2, describe the GRAVITY data in Sect. 3 and apply $G^R$ to it in Sect. 4. The images reveal the rich structure of sources around the SMBH and a previously unknown faint star, moving to the west at high angular velocity. In Sect. 5, we first compare the results of the new algorithm to image reconstruction with CLEAN, before we discuss the properties and implications of the newly detected star. Finally, we present our conclusions in Sect. 6.

## 2. Imaging method

The backbone of our method is a Bayesian interpretation of the imaging process, in which the posterior probability for a certain image, $I$, given the data, $d$, is proportional to

$$\mathcal{P}(I|d) \propto \mathcal{P}(d|I)\,\mathcal{P}(I)\,, \qquad (1)$$

that is, the prior times the likelihood. The proportionality arises because we have omitted the notoriously difficult-to-calculate evidence term $\mathcal{P}(d)$, which is insensitive to the particular image realization (we come back to this choice in Sect. 2.5). If the posterior distribution is known, we can easily find the most likely

---

[1] The $G^R$ source code is available at `https://gitlab.mpcdf.mpg.de/gravity/gr_public.git`

image from it or we can compute the expected image and its variance from the first two posterior moments. In practice, however, the posterior is a very high-dimensional scalar function (each image pixel contributes one dimension), and we cannot evaluate its moments analytically. It is then necessary to resort to numerical methods, such as descent minimization to find a (local) maximum or to sampling techniques.

In this section, we introduce all the necessary ingredients to implement the Bayesian inference scheme in practice. We start with the prior (Sect. 2.1), which is formulated as a generative model such that random samples can be drawn from it. Each sample constitutes a possible realization of the image before knowledge of the data. It is then processed by the instrumental response function (Sect. 2.2) to arrive at a prediction of the data. Some time-varying instrumental effects, such as coherence loss over a baseline, cannot be described to sufficient accuracy by a deterministic model, and we account for them in a self-calibration approach (Sect. 2.3). The agreement between predicted and actual data is then quantified by the likelihood (Sect. 2.4). With this, all the components are in place to compute the posterior of a certain sample. Finally, the exploration of the posterior is performed with the MGVI algorithm (Sect. 2.5).

### 2.1. Prior model for the Galactic Center

A major challenge in imaging the GC is the variability of Sgr A*. During very bright epochs, it can outshine ambient stars and the signal more strongly resembles the expectation for a single point source. When Sgr A* is fainter, on the other hand, signatures of the stars in its vicinity become very apparent. Continuous alteration of the flux, between the two extreme cases, prevents the naive combination of exposures over a longer observing period. We account for it by describing Sgr A* as a time-variable point source that we superimpose on the actual, static image.

In our prior model, the position of Sgr A*, $s_{\text{SgrA}}$, follows a Gaussian distribution with user-defined mean and variance. For the flux $I_{\text{SgrA}}(t, p)$, we impose a log-normal distribution, and account for the variability by inferring an independent flux value for each exposure $t$. While it would, in principle, be possible to explicitly model the temporal correlation, namely, along the lines of Arras et al. (2020), this further increases the size and complexity of the parameter space, introduces additional hyperparameters, and makes the sampling more demanding. We therefore leave the exploration of this option to future studies. Furthermore, the emission from Sgr A* is known to be polarized, with the polarization also varying on short timescales (Trippe et al. 2007; Zamaninasab et al. 2010; Shahzamanian et al. 2015; Gravity Collaboration et al. 2018b, 2020d). We thus also allow for Sgr A* to have differing, independent fluxes in each of the two polarization states observed by GRAVITY and which we have labeled as $p$.

The actual image, $I_{\text{Img}}(s)$, then contains the stars in the vicinity of Sgr A*, where $s$ refers to the angular direction relative to the phase center. It must be large enough to cover the full FOV of the GRAVITY fibers, which have a full width half maximum (FWHM) of about 65 mas. On the other hand, the pixel scale $\delta\alpha$ needs to be sufficiently small to describe the highest spatial frequencies or smallest scales observed by GRAVITY. The latter requirement is the Nyquist–Shannon theorem to avoid spectral aliasing and imposes $\delta\alpha \leq \lambda/(2b_{\text{max}})$ (Thiébaut 2008). With the longest baseline of 130 m and the smallest used spectral channel at $\lambda_{\text{min}} = 2.11\,\mu\text{m}$, this requires $\delta\alpha \lesssim 1.7$ mas. By using a grid with $256^2$ pixels of 0.8 mas size, we can meet both requirements simultaneously at manageable computational costs.

To GRAVITY, stars in the GC appear as unresolved point sources. Consequently, we assume that all pixels in the image are statistically independent. We note that this statement applies to the sky prior model and that correlations between nearby pixels introduced by the instrument's finite spatial resolution are described by the response function (Sect. 2.2). Furthermore, we expect to see only $O(1)$ stars in the image such that the vast majority of pixels will be dark. We describe this situation by imposing an inverse-Gamma prior on the brightness of each pixel:

$$\mathcal{P}\left(I_{\text{Img}}\right) = \prod_i^{N_{\text{pix}}} \frac{q^\alpha}{\Gamma(\alpha)} I_{\text{Img}}(i)^{-\alpha-1} \exp\left[\frac{-q}{I_{\text{Img}}(i)}\right]. \qquad (2)$$

Here, the index $i$ labels all pixels in the image, $\Gamma$ is the Gamma function, and $q$ and $\alpha$ are two parameters that need to be specified before applying the code. They determine the a priori probability of encountering a bright star (where larger $\alpha$ implies a smaller probability) and allow for the mode of the distribution, $q/(\alpha+1)$, to be set to some background level.

As discussed at length in the following section, the primary observables in optical and near-IR interferometry are the complex visibilities (cf. Eq. 3), given by the ratio between the coherent flux of a baseline and the flux on each of its two telescopes. Consequently, the data is only sensitive to the relative brightness of all sources. The maximum possible value for the visibility amplitude equals unity and decreases in the presence of a homogeneous background. With this in mind, we set $q = 10^{-4}$ and $\alpha = 1.5$. In a typical GC observation, the instrumental visibilities are on the order of 0.5, such that this setup implies $O(1)$ point sources. The probability of encountering a unit point source in the image of faint sources, on the other hand, is $\sim 5\%$. We further chose zero as the mean and unity as the variance for the Sgr A* log-flux prior, whereby the large flexibility of the log-normal distribution can easily accommodate changes by an order of magnitude.

In many observations, there are bright stars within the FOV around Sgr A*, whose positions can be adequately predicted from their known orbits. We incorporate them into our model by allowing for additional point sources with a Gaussian prior on their position and a log-normal flux prior. In contrast to Sgr A*, we assume the flux to be constant in time and the same for both polarization states. In principle, these stars could also be attributed to the image; modeling them explicitly helps to mitigate pixelization errors and improves the convergence.

Finally, the model describes the spectral distribution of Sgr A* and the stars as two power laws with different spectral indices. These follow a Gaussian prior whose mean we set to $-1$ and $+2$ for Sgr A* (Genzel et al. 2010; Witzel et al. 2018) and the stars respectively. As variance, we assume 3 for both the stars and Sgr A*. A mathematical description of the sky model is provided in Eq. (A.3), and the priors on each of its components are noted down explicitly in Eq. (B.2) to Eq. (B.7).

### 2.2. Instrumental response function

The primary observable of GRAVITY is the complex visibility, measured as the coherent flux over a baseline, $\boldsymbol{b}$, divided by the flux on each telescope forming that baseline. In an idealized setting, it is related to the sky flux distribution, $I(s, \lambda)$, by

$$v(\boldsymbol{b}/\lambda) = \frac{\int_{\text{FOV}} \text{d}s\, I(s, \lambda)\, e^{-2\pi \boldsymbol{b}\cdot s/\lambda}}{\int_{\text{FOV}} \text{d}s\, I(s, \lambda)}, \qquad (3)$$

where $\lambda$ is the wavelength. In reality, instrumental effects render the response equation more complicated. We discuss the modifications qualitatively in the following and the generalized expression is given in Eq. (A.1).

The relevance of the denominator in Eq. (3) is most intuitively understood by considering a homogeneous background flux, which averages out in the numerator term for $b \neq 0$ but affects the denominator. We therefore implemented the Fourier transform and the normalization term explicitly in the response function.

Equation (3) constitutes a simplification in that not all sources are coupled into the fiber equally. GRAVITY uses single mode fibers, which have a Gaussian acceptance profile, to transport the light from the telescope to the beam combiner (Gravity Collaboration et al. 2017). The coupling of light into the instrument is further affected by diffraction (Shaklan & Roddier 1988; Guyon 2002) and residual tip-tilt jitter from the AO system (Perrin & Woillez 2019). Both effects lead to a wider FOV than implied by the fiber profile alone. Furthermore, there are static optical aberrations along the GRAVITY light path which impact also the phase of the transmitted light (Gravity Collaboration et al. 2021b). These aberrations are different for each of the four telescopes.

Taken together, fiber damping and optical aberrations act as a position dependent phase screen which multiplies the image inside the Fourier transform. With the measurement of these so-called phase- and amplitude-maps in Gravity Collaboration et al. (2021b), we are able to provide a full model of the optical aberrations and fiber damping, which we include in the implementation of the response function.

GC observations with GRAVITY are carried out in low spectral-resolution mode and the effect of the spectral bandwidth on the measurement needs to be taken into account. This is known as bandwidth smearing and it leads to a further modification of Eq. (3), where the numerator and denominator are independently integrated over the bandpass, namely, $\int d\lambda P(\lambda)$. In the most simple case of flat spectral distributions, bandwidth smearing multiplies the visibility amplitudes by a sinc function. For a more realistic scenario in which the source's spectral distribution is modeled as a power law, bandwidth smearing is described by a generalized complex gamma function and, thus, it also affects the visibility phases. Either way, the effect on the measured visibilities increases with the source distance from the image center.

We account for the finite bandpass size by evaluating the Fourier transform in Eq. (3) at multiple values of $\lambda$, spread equally over a top-hat bandpass and taking the average. The complete response equation and details about its efficient implementation are provided in Appendix A.

### 2.3. Self-calibration

There are time-variable instrumental effects that our response function does not capture, such as coherence loss on a baseline or telescope-dependent phase errors, for instance, from atmospheric conditions or instrumental systematics. The former leads to a reduction of the measured visibility amplitude, while the latter impacts the visibility phase. If unaccounted for, either of these effects can significantly degrade the imaging solution.

We resolve telescope-dependent phase errors by working with closure phases (Rogers et al. 1974),

$$\phi_{i,j,k} = \arg\left(v_{i,j}\, v_{j,k}\, v_{k,i}\right), \tag{4}$$

which are formed over a triangle of telescopes $i$, $j$, and $k$, such that telescope-based errors are canceled out. The VLTI consists of four telescopes, implying that six visibility phases and four closure phases can be measured per spectral channel.

GRAVITY carries out a visibility measurement for each baseline. With regard to the amplitudes, we thus adopted a self-calibration approach in which we infer an independent scaling factor, $C(t, b)$, for each exposure and baseline, $b$. This factor joins the list of free model parameters in Appendix B. We impose a Gaussian prior on the scaling factor with unit mean and standard deviation 0.1, as given by Eq. (B.7).

Visibility amplitudes and closure phases are both translation-invariant, that is to say that in the absence of absolute phase information, any global shift of the image will not affect the posterior. Phase- and amplitude maps as well as bandwidth smearing partly break the degeneracy, but these instrumental effects are not sufficiently strong to reliably center the image, particularly if the brightest sources are located close to the image center. Instead, we fix the position of one point source, usually Sgr A*, which we refer to as as the anchor for our images.

### 2.4. Likelihood

In the limit of sufficiently high signal-to-noise ratios (S/N), that is S/N $\gtrsim 2 - 5$, the noise properties of visibility amplitudes and closure phases are well approximated by a Gaussian distribution (Blackburn et al. 2020). This condition is well satisfied for typical GC observations with GRAVITY.

However, only three of the four closure triangles available to GRAVITY are statistically independent. In principle, we need to account for this by selecting a reduced, non-redundant closure set whose cross-correlations are captured by a non-diagonal covariance. In practice, and in the limit of equal S/N on all telescopes, equivalent likelihood contours can be obtained from the full closure set without accounting for cross-correlations by upscaling the error bars with a redundancy factor $N/3$, where $N$ is the number of telescopes (Blackburn et al. 2020). In this work, we adopt the latter formulation.

Forward modeling approaches are rather sensitive to underestimated error bars. In the case of our imaging code, we observe that overfitting the data can introduce spurious sources in the images. To avoid any potential bias, we applied a conservative $O(1)$ factor to the error bars in the first imaging iteration. This factor may differ for closures and amplitudes but it is common to all frames and baselines. We then checked the residuals and adjust the error scaling for subsequent runs, aiming for a reduced $\chi^2$ of $0.95-1.0$ for visibility amplitudes. To realize the aforementioned redundancy factor, the targeted reduced $\chi^2$ for the closure phases is smaller than the noise expectation from Gaussian statistics and in the range of $0.7 - 0.75$.

### 2.5. Inference strategy

The goal of the inference is to explore the posterior distribution in Eq. (1) around its maximum over a very high dimensional parameter space. Indeed, for a data set consisting of $N_{\mathrm{exp}}$ individual exposures and considering $N_{\mathrm{PS}}$ static point sources in the prior

model, the dimensionality is

$$
\begin{aligned}
d = 256^2 & \quad \text{image of faint sources} \\
+ 2 \times N_{\text{exp}} & \quad \text{Sgr A* light curves} \\
+ 2 & \quad \text{Sgr A* position} \\
+ 3 \times N_{\text{PS}} & \quad \text{point sources position and flux} \\
+ 6 \times N_{\text{exp}} & \quad \text{amplitude self-calibration} \\
+ 2 & \quad \text{spectral indices} \quad (5) \\
\sim 7 \times 10^4 \, . &
\end{aligned}
$$

Here, the factor of two in the light curves arises from the two polarization states observed by GRAVITY and there are six baselines available. Dealing with such high-dimensional distributions is notoriously difficult and computationally expensive.

The MGVI algorithm (Knollmüller & Enßlin 2019) used in this study searches for a multivariate Gaussian distribution $\mathcal{G}\left(\xi | \bar{\xi}, \Xi\right)$ that best approximates the full posterior. Here, $\xi$ are standardized coordinates, thus we mapped each degree of freedom to an auxiliary parameter or excitation $\xi$, whose prior is given by a unit Gaussian with zero mean. The inference targets the excitations rather than the physical quantities, however, the two are uniquely related by a mapping (given in Appendix B). As an advantage, this setup makes it fast and easy to draw samples from the prior.

The mean $\bar{\xi}$ and covariance $\Xi$ of the posterior approximation are found in an iterative procedure. Starting at some initial position $\bar{\xi}_i$ ($i = 0$), MGVI uses a generalization of the Fisher metric (cf. Eq. B.11) to approximate the covariance. This is a $d \times d$ matrix, and by allowing for non-zero off-diagonal elements, MGVI is able to capture cross-correlations between individual model parameters. Since the explicit storage of $d^2$ matrix elements becomes prohibitive for large inference problems, an implicit representation in the form of a numerical operator is used internally.

In the next step, the mean of the approximate distribution $\bar{\xi}$ is updated with the aim to increase the overlap between approximate and true posterior $\mathcal{P}(\xi|d)$. This overlap is quantified by the Kullback-Leibler divergence (KL), such that updating $\bar{\xi}$ amounts to minimizing the KL, that is,

$$
\bar{\xi}_{i+1} = \min_{\bar{\xi}} \int \text{d}\xi \, \mathcal{G}\left(\xi | \bar{\xi}, \Xi_i\right) \ln \left[ \frac{\mathcal{G}\left(\xi | \bar{\xi}, \Xi_i\right)}{\mathcal{P}\left(\xi | d\right)} \right] . \quad (6)
$$

The minimum is insensitive to any multiplicative factors applied to the true posterior as long as they are independent of $\xi$, thus, the omission of the evidence term in Eq. (1), in fact, does not affect our analysis. To estimate the expectation value in the KL numerically, MGVI draws samples from the approximate posterior distribution and replaces the integral in Eq. (6) by the sample mean. Once $\bar{\xi}$ has been updated, MGVI computes the covariance at the new position and henceforth alternates between the $\Xi$ and $\bar{\xi}$ determination.

After the final iteration, the primary product is a set of samples distributed according to the approximate posterior distribution. These samples can then be used to compute expectation values and their standard deviations, which we report as our main results.

We note that MGVI leaves it to the user to set the number of samples and minimization steps at each iteration as well as the total number of iterations. A good practice is to start with few samples and steps, then increase both quantities as the inference approaches the posterior maximum. Details about the scheme used for this work are given in Appendix C, where we also specify the minimizers and iteration controllers used. For the initial

position from which to start the inference, we align all parameters with the mode of their respective priors.

Two obstacles in the minimization procedure demand special attention, namely: the convergence of our results and the possibility of multi-modal posterior distributions. In case of the latter, the Gaussian approximation obtained from MGVI captures one typical posterior mode. Exploring multi-modal distributions in such high-dimensions is genuinely a very difficult, computationally expensive problem for which there currently exists no standard practice. To maintain some handle on the multi-modality of the posterior and to judge how well the algorithm has converged, we exploit the inherent stochasticity of MGVI that arises when the KL is estimated from random samples. As our results below indicate, changing the random seed from which samples are drawn can nudge the algorithm to explore different posterior modes. In addition, poorly converged runs – which are typically more noisy and might contain over-fitted sources or might fail to detect a faint source – also occur just for individual random seeds, such that they can be judged and eliminated by comparing results from multiple random seeds.

For a typical data set, we performed ten independent imaging runs with different random seeds. This number is large enough to capture the dominant modes of the posterior but too low to determine their relative weights reliably. Instead, we used the fact that we have multiple data sets available and judged the images by their consistency over the full 2021 observing period.

## 3. Data

For this study, we considered GC observations with GRAVITY from 2021. They are separated into four epochs taking place at the end of March, May, June, and July, respectively. The data were obtained with GRAVITY's low spectral-resolution, split-polarization mode. Each exposure consists of 32 individual frames with 10 s integration time. Most of them are centered on Sgr A*, and the Sgr A* observing blocks are bracketed by S2 and R2 pointings that are required for calibration and to monitor the S2 orbit. The data were reduced by the GRAVITY standard pipeline and calibrated with a single, carefully chosen S2 exposure from the same night.

We used the data to obtain night-wise images of Sgr A* and its vicinity. To have sufficient $(u, v)$-coverage, we required $O(10)$ exposures at least and selected those nights with a larger number of Sgr A* pointings available. The selection is summarized in Table 1.

The Sgr A*-centered images from March to June revealed a faint object, moving to the west at high angular velocity. To track this new star and to also scan a wider field, we performed a series of pointings in July, where we offset the fiber from Sgr A*. Further, to the south west of Sgr A* there is S38 with a separation too large to be detected in the Sgr A*-centered images. For the July observing run, there are a sufficiently large number of S38-centered exposures available to attempt reconstructing an image from them. Just as with the Sgr A*-centered exposures, the data were reduced by the standard pipeline and calibrated with a single S2 file. Finally, we also considered S2-pointings from May, which were calibrated with R2. The so-called "mosaicing data set" is summarized in Table 2.

## 4. Results

We apply the $\mathrm{G}^{\mathrm{R}}$ imaging algorithm (introduced in Sect. 2) to obtain night-wise images from the data summarized in Tables 1
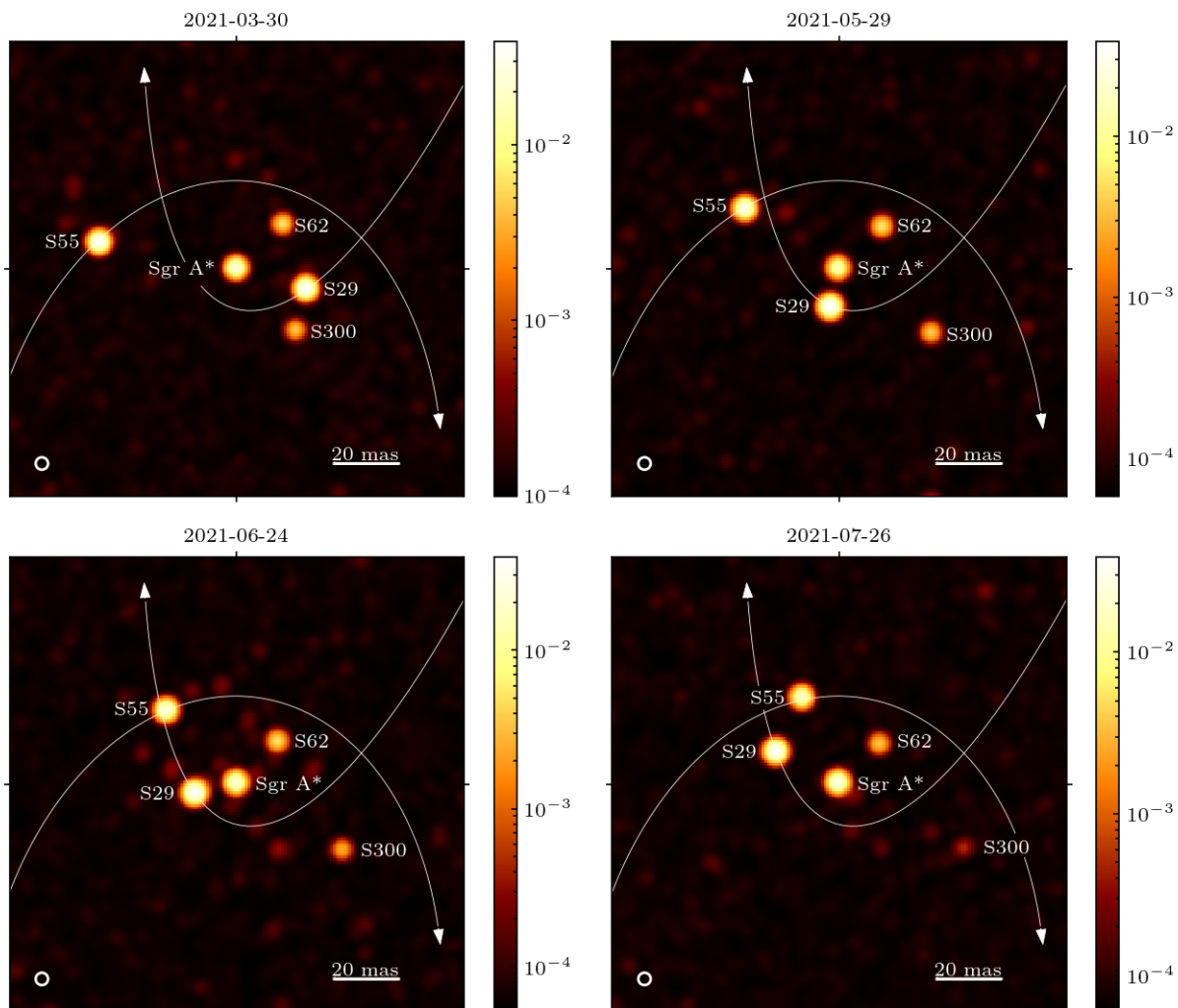
**Fig. 1.** Images for one night per observing epoch in 2021, where north is up, east to the left, and the flux normalized to S29. For each night, we performed ten imaging runs with varying random seeds and then picked the cleanest result with the highest degree of consistency over the full year. Images are computed as mean over all samples in the chosen run, centered on Sgr A*, and smoothed with a Gaussian of 1.6 mas standard deviation, whose FWHM is indicated in the bottom left corner. Sgr A* varies in flux, and we here show its mean brightness over all exposures and polarization states. On each image, we overdraw the orbits of S29 and S55 and the labels of all identified sources to give a better orientation. We note that due to the large dynamical range and the logarithmic color scaling, sources appear more widespread.

and 2. Here, we first focus on observations with Sgr A* at the center and come back to the mosaicing data set later on.

The final output of the MGVI algorithm is a set of samples drawn from the approximate posterior. Our best image is computed as the mean over these samples and we add the flux of all additional point sources to the appropriate pixels. Because of the normalization term in the instrumental response function (cf. Eq. 3), the likelihood is insensitive to a global scaling of the flux. In a post-processing step, we therefore normalize our images to the flux of S29 ($m_K$ (S29) $\simeq$ 16.6, Gravity Collaboration et al. 2021a), which is the brightest static source in the field. The implementation of the response function presented in Appendix A explicitly takes into account fiber damping, which reduces the flux transmission for off-center sources; thereby, our images are automatically corrected for this effect. Finally, we smoothed the images with a Gaussian of 1.6 mas standard deviation. This corresponds to the typical size of the CLEAN beam, that is, a Gaussian fit to the central part of the dirty beam pattern.

### 4.1. Detection of S29 and S55 within 50 mas of Sgr A*

In 2021, there are two relatively bright stars in the FOV around Sgr A*, which are S29 with a K-band magnitude of $m_K$ (S29) $\simeq$ 16.6 (Gravity Collaboration et al. 2021a) and S55 with $m_K$ (S55) $\simeq$ 17.6 (Gillessen et al. 2017). In particular the cross-identification of S29 in earlier AO-based NACO images has been subject of debate recently (Peißker et al. 2021). We discuss this in Appendix D.

The orbital motions of S29 and S55 around the black hole are overdrawn on the images in Fig. 1. In a couple of test runs, we find that our algorithm is easily capable of detecting each of them. The detection of S29 and S55 in the 2021 Sgr A*-centered exposures also is a prerequisite and starting point for astrometric fitting to extract high-accuracy stellar positions.

In what follows, we model S29 and S55 as point sources, superimposed on the image as described in Sect. 2.1. Each of the sources has a Gaussian prior on its position, whose mean we obtain either from orbit predictions or from the pixel position in a pre-imaging run. To avoid any bias arising from the latter pro-
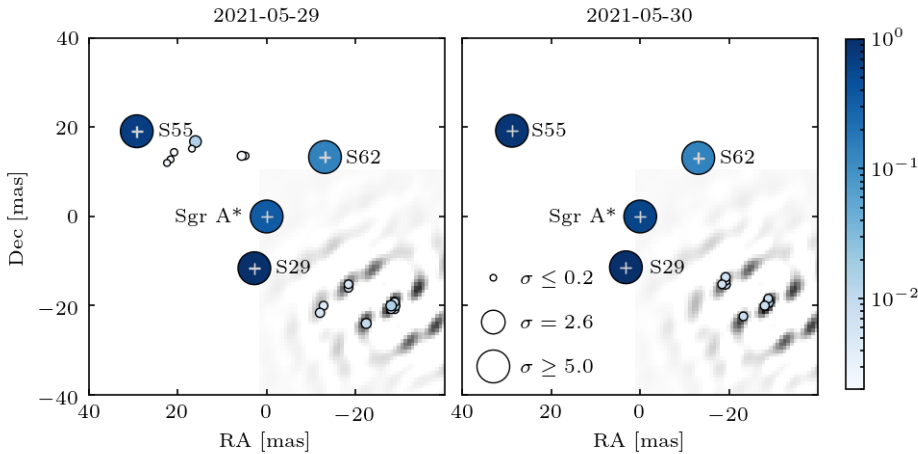
**Fig. 2.** Combined view on all imaging runs from a single night (see Appendix E for details on the illustration). The symbol color indicates the flux of a source candidate, normalized to S29, while the symbol size represents its significance. The Sgr A* flux depicted here has to be multiplied by the light curve at each exposure to arrive at the true flux ratio. For stars that are modeled as a point source, the position uncertainty is indicated in gray. Sources are shown superimposed on the dirty beam pattern of the imaging night, which we have centered at the position of S300.

**Table 1.** GC observing nights selected for Sgr A* imaging.

| Date | Available exposures | $N_{exp}$ | $\langle\sigma_\rho\rangle$ | $\langle\sigma_\phi\rangle$ |
|---|---|---|---|---|
| 2021-03-30 | 13 | 13 | $4.5 \times 10^{-2}$ | 33.2° |
| 2021-05-22 | 20 | 19 | $3.3 \times 10^{-2}$ | 10.5° |
| 2021-05-29 | 22 | 21 | $3.5 \times 10^{-2}$ | 19.3° |
| 2021-05-30 | 20 | 20 | $3.7 \times 10^{-2}$ | 24.3° |
| 2021-06-24 | 32 | 29 | $3.6 \times 10^{-2}$ | 20.7° |
| 2021-07-25 | 28 | 22 | $3.1 \times 10^{-2}$ | 15.3° |
| 2021-07-26 | 20 | 20 | $3.6 \times 10^{-2}$ | 19.6° |
| 2021-07-27 | 20 | 19 | $2.8 \times 10^{-2}$ | 10.3° |

**Notes.** Each exposure amounts to a 320 s integration time. For some nights, we deselected exposures which were affected by bad weather or instrumental problems, such that the number of exposures used in the imaging, $N_{exp}$, is smaller than the total number of exposures available. The last two columns give the mean standard deviation of the closure phases and amplitudes. If a scaling factor was applied to the error bars, this is included.

cedure, we chose a deliberately large standard deviation for the sources' position priors of 2.4 mas in RA and Dec. This corresponds to three pixels in our image and is larger than the beam width in any of the observations considered. Since Sgr A* is fixed at the image center to break the translation invariance inherent to closure phases, mapping the stars' separation vector to an image position is straightforward.

### 4.2. Detection of S62

In Gravity Collaboration et al. (2021a), CLEAN images revealed a 18.9 K-band magnitude star which slowly approaches Sgr A* and was identified as S62. Extrapolating the motion observed in 2019, we expect S62 at RA = $(-14.1 \pm 0.4)$ mas, Dec = $(13.6 \pm 0.8)$ mas in March 2021. By July, it should move to RA = $(-13.2 \pm 0.4)$ mas, Dec = $(12.3 \pm 0.9)$ mas.

An important test for the new imaging method is whether it is also able to detect S62. Indeed, for all nights we infer flux at the expected position. This detection is very robust, namely, $> 5\sigma$ for almost all random seeds and even holds if the error bars are moderately over-scaled. To determine the S62 coordinates beyond the accuracy of the pixel size, we then include it into the set of point sources inferred on top of the image and perform ten imaging runs with varying random seeds for each night. The results are consistent between all runs of an individual night. We

are thus able to combine the samples from all ten runs into an estimate on the mean S62 position and its variance. This is summarized in Table 3 and it does match the prediction from Gravity Collaboration et al. (2021a) very well.

Furthermore, we use the positions in Tab 3 to provide an updated on the motion of S62. From a linear fit that also considers the results from 2019 observations (Gravity Collaboration et al. 2021a), we obtained the following for the relative velocity w.r.t. Sgr A*:

$$v_{RA} = (2.97 \pm 0.05) \text{ mas/yr},$$
$$v_{Dec} = (-3.58 \pm 0.09) \text{ mas/yr}.$$

The slow linear motion of S62, observed with GRAVITY consistently in 2019 and 2021, does not fit a star with a 9.9 year orbital period as reported in Peißker et al. (2020). In Appendix D, we explain in detail the cross-identification of all further sources in the FOV, S29, and S55, namely, between GRAVITY and earlier AO-based images. Thus, the GRAVITY images do not support the existence of a star that orbits Sgr A* with a 9.9 year period.

### 4.3. Discovery of S300, a faint fast-moving star

The main goal of our imaging analysis is to search the vicinity of Sgr A* for faint, as-yet-unknown stars. To this end, we used the same ten imaging runs per night from which we measured the S62 position (detailed in Sect. 4.2). A representative image for each epoch is shown in Fig. 1. In Appendix E, we provide a complementary view of the imaging results for all nights listed in Table 1, which accounts for the statistical nature of our algorithm.

In addition to the four expected sources – Sgr A*, S29, S55, and S62 – our images contain a fifth object that is fainter than all aforementioned stars. It moves to the west with a high angular velocity, so that the change in position can be recognized very clearly between individual months. In the following, we discuss this detection for each epoch separately.

#### 4.3.1. March 2021

Due to the limited observability of the GC in March, this is the data set with the sparsest $(u, v)$-coverage. However, it is also the epoch where the new source is closest to the center of the FOV and thus it is the least affected by fiber damping. Of the ten imaging runs we performed, five detected the new source as a single

**Table 2.** Overview of the mosaicing data set.

| | | | Pointing offset w.r.t | | |
| Name | Date | # of exposures | Sgr A* (RA, Dec [mas]) | preset sources | anchor |
|---|---|---|---|---|---|
| S2 | 2021-05-29 & 30 | 13 | 24.8, 142.4 | S2 | S2 |
| NW | 2021-07-29 | 8 | −45.0, 45.0 | Sgr A*, S29, S55, S42 | S42 |
| SE | 2021-07-29 | 8 | 45.0, −45.0 | Sgr A*, S29, S55 | Sgr A* |
| mid | 2021-07-29 | 7 | −24.8, −31.4 | Sgr A*, S29, S55 | Sgr A* |
| S38 | 2021-07-25 & 26 | 8 | −38.6, −76.8 | S38 | S38 |

**Notes.** The second to last column lists all sources which are modeled as point sources with a Gaussian position prior. To break the translation invariance inherent to closure phases and visibility amplitudes, we fix the location of one bright point source, which we call the anchor for that image.

**Table 3.** Separation between S62 and Sgr A* obtained from imaging runs in which S62 was modeled as point source with a Gaussian position prior.

| Epoch | RA [mas] | Dec [mas] |
|---|---|---|
| 2021.2453 | −13.85 ± 0.11 | 14.00 ± 0.10 |
| 2021.3902 | −13.03 ± 0.05 | 13.12 ± 0.06 |
| 2021.4093 | −13.12 ± 0.06 | 13.33 ± 0.07 |
| 2021.4120 | −13.05 ± 0.06 | 13.13 ± 0.08 |
| 2021.4803 | −12.78 ± 0.04 | 12.84 ± 0.05 |
| 2021.5649 | −12.60 ± 0.04 | 12.67 ± 0.05 |
| 2021.5657 | −12.74 ± 0.05 | 12.64 ± 0.07 |
| 2021.5703 | −12.71 ± 0.09 | 12.57 ± 0.14 |

**Notes.** The epoch is computed as mean over all exposures used for the image. For each night we performed ten $G^R$ runs with varying random seeds and combined the samples from all runs into an estimate of the S62 position and its standard deviation. We note that the standard deviation only accounts for the statistical position uncertainty, but not for any systematic error.



**Fig. 3.** Light curves inferred with the $G^R$-code for the May observing nights. Light lines in gray and red indicate the individual samples and dark lines show the sample mean. Here, we combined all samples from all imaging runs of a particular night. The time of the individual exposures is indicated by vertical lines. We note that S2 is about 11 times as bright as S29.

bright pixel with high significance ($> 5\sigma$). Its location on the grid can vary by one pixel between runs. Two further runs infer flux at the same position, but smeared out over multiple neighboring pixels.
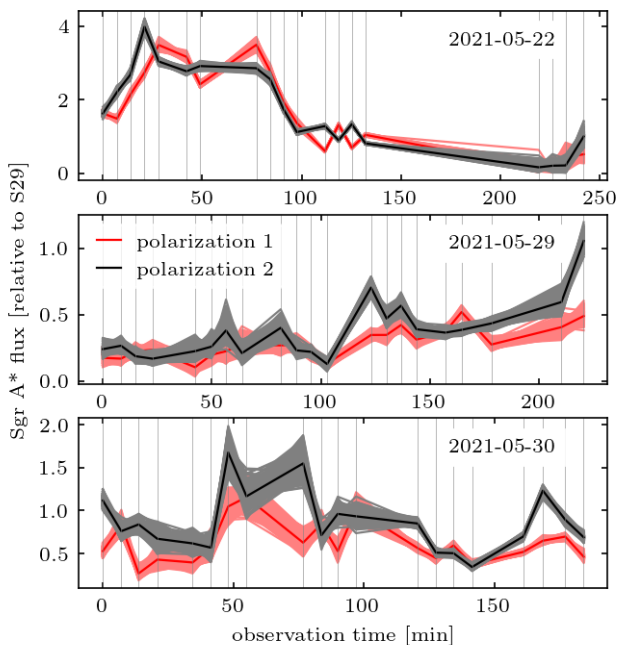
### 4.3.2. May 2021

In Sect. 2.5, we mention the possibility of multi-modal posterior distributions. Indeed, the only data set where we have clear signs of such an issue are the three Sgr A*-centered images from May 2021.

All ten imaging runs for May 29 infer a single bright pixel to the south west of Sgr A* at a high significance. In three instances, however, the location of this pixel is shifted towards the image center. A similar situation arises for May 30. In four instances, the location of the bright pixel coincides with the position found for the previous night, otherwise it is shifted inwards.

We illustrate this situation in Fig. 2, where we have combined the samples from all ten imaging runs into a single figure for each night. Even though the new source is detected with a high significance in individual $G^R$-runs, the fact that its location varies between runs makes the overall significance estimate decrease in Fig. 2 (cf. Appendix E). The sources detected in the image are superimposed on the dirty beam pattern of the respective nights to illustrate the reason behind the observed multimodality. This shows that the inward-shifted detections of the new source correspond to side-maxima or side-lobes of the dirty beam pattern.
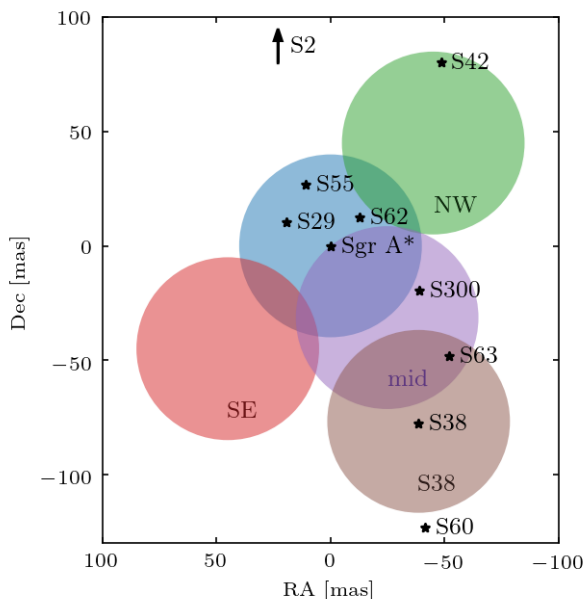
On May 29, the algorithm apparently is somewhat more successful at disentangling the true source position and its sidelobes than on May 30. There are two possible reasons for this, the first being that the data set for May 29 contains one more exposure and on average exhibits smaller error bars than that of May 30, as Table 1 indicates. Secondly, the light curves that we obtain as part of the inference are shown in Fig. 3 for all May nights. They disclose that Sgr A* was slightly brighter on May 30 than on the 29th. When the complex visibilities are dominated by the central source more strongly, the detection of the faint source further out in the field becomes more difficult.

The images that we infer for May 22 also are consistent with that line of reasoning. As Fig. 3 shows, Sgr A* went through a moderate flare in the beginning of the night. While all imaging runs for May 22 clearly exhibit some flux to the south west of Sgr A*, no consistent source position can be identified from the comparison of multiple runs (cf. Fig. E.2).

**Fig. 4.** Summary of pointings in the mosaicing data set in the relation to the positions of all detectable stars during the July observing run. For completeness, we further include the Sgr A*-centered exposures (blue). Colored circles indicate the individual pointings and have diameters of 40 mas, which is half the extent of the images in Fig. 5.

### 4.3.3. June 2021

For the large June imaging data set, all ten imaging runs exhibit flux at the same location to the south west of Sgr A*. In five instances, this flux is detected in a single pixel with high significance ($> 5\sigma$). Its position scatters by at most a pixel. In the other five instances, the flux is spread out over several connected pixels, which exhibit larger flux variations between the individual samples.

Since the source has moved further away from the field center and is more strongly affected by fiber damping than it was in May, the question arises why no similar multi-modality of the posterior is observed. On June 24, we have a much better $(u, v)$-coverage and the dirty beam pattern becomes considerably smoother without the isolated, strongly peaked side-maxima that induce the misplacement in the May images.

### 4.3.4. July 2021

We are able to detect the new source only in one of the July observing nights which is July 26. Even so, it is found in five of the ten imaging runs at a low significance smeared out over multiple pixels. Correspondingly, the source appears fainter in the final image of Fig. 1. Since the new star is most strongly affected by fiber damping in July, this loss of sensitivity does not come as a surprise. We discuss it further in Sect. 4.5.

We summarize the source position inferred for each night in Table 4, along with its flux relative to S29. For the latter, we selected all runs in which the new source is found as a single bright pixel in the correct location and combined all samples in these runs. That is, we exclude runs from the flux estimate that failed to identify the new source, where the new source was misplaced at a side-lobe, or where its flux was smeared out over multiple pixels.

The position and motion of the new star matches none of the known S-stars (Gillessen et al. 2017) and we conclude that we

**Table 4.** Position of S300, the newly detected star, with respect to Sgr A*.

| Epoch | RA [mas] | Dec [mas] | flux/S29 |
|---|---|---|---|
| 2021.2453 | −18.0 ± 0.8 | −19.6 ± 0.8 | 0.11 ± 0.01 |
| 2021.4093 | −28.4 ± 0.8 | −20.4 ± 0.8 | 0.09 ± 0.01 |
| 2021.4120 | −28.4 ± 0.8 | −18.8 ± 0.8 | 0.08 ± 0.01 |
| 2021.4803 | −33.4 ± 0.8 | −20.4 ± 0.8 | 0.05 ± 0.02 |
| 2021.5657 | −39.6 ± 0.8 | −19.6 ± 0.8 | – |

**Notes.** The flux ratio relative to S29, is already corrected for fiber damping and the epoch is computed as mean over all exposures used in the imaging.

have detected a new star, which we refer to as S300. We further discuss the possibilities for its nature in Sect. 5.2.

### 4.4. Images for the mosaicing data set

To produce images for the mosaicing data set, we used the exact same strategy as for the Sgr A*-centered exposures. The pointing directions and the known stars in the field are summarized in Fig 4. In Table 2, we list all stars that we model as point sources with a Gaussian position prior and the anchor, whose positional variance we set to zero in order to break the translation invariance inherent to closure phases and visibility amplitudes (cf. Sect. 2.3).

In contrast to the Sgr A*-centered images and with exception of the S2 pointings, we now have fewer than ten exposures available for each individual image. We thus expect (and also go on to find) that the solutions become more noisy and that there is a greater variability between the ten runs which we perform for each data set. The images obtained for the mid, NW, and S38 pointings are shown in Fig. 5, and the full statistical view over all results is provided in Appendix E and in Fig. E.3.

The mid pointings were specifically designed to provide an additional test of the S300 detection. The fiber position is such that S300 is close to the center of the field, but Sgr A*, S63, and S38 can also be observed within the same pointing; Fig. 5 very clearly shows all the expected sources. The new source, S300, is detected with high significance ($> 5\sigma$) in each of the imaging runs, its position varies by at most one pixel in RA and two pixels in Dec. Combining all mid-pointing imaging runs, we obtain:

$$\text{RA} = (-38.8 \pm 0.8)\,\text{mas}\,,$$
$$\text{Dec} = (-19.4 \pm 1.6)\,\text{mas}\,,$$
$$\text{flux/S38} = 0.15 \pm 0.10\,,$$

which is fully consistent with the position obtained from Sgr A*-centered images in Table 4.

An important difference to the Sgr A*-centered images, in particular, for the mid and S38 pointings, is that we did not preset all known sources in the field. In the former case, it was left to our algorithm to detect S300, S63, and S38; in the latter case, S63 and S60. In this context, the images in Fig. 5 also demonstrate how $G^R$ is able to orient in a field with limited prior knowledge about the overall source structure. Apart from the stars that are sketched in Fig. 4 – namely Sgr A*, S2, S29, S38, S42, S55, S60, S63, and S300 – we do not detect any other objects.

The detection of all the aforementioned sources in the images also serves as a starting point for astrometric fitting to obtain high accuracy source positions (see Gravity Collaboration et al. 2020b, for details on the fitting procedure), which, in turn, allows us to determine the stars' orbits. We give an update on the orbital elements of all stars detected in the 2021 images in Table 5.
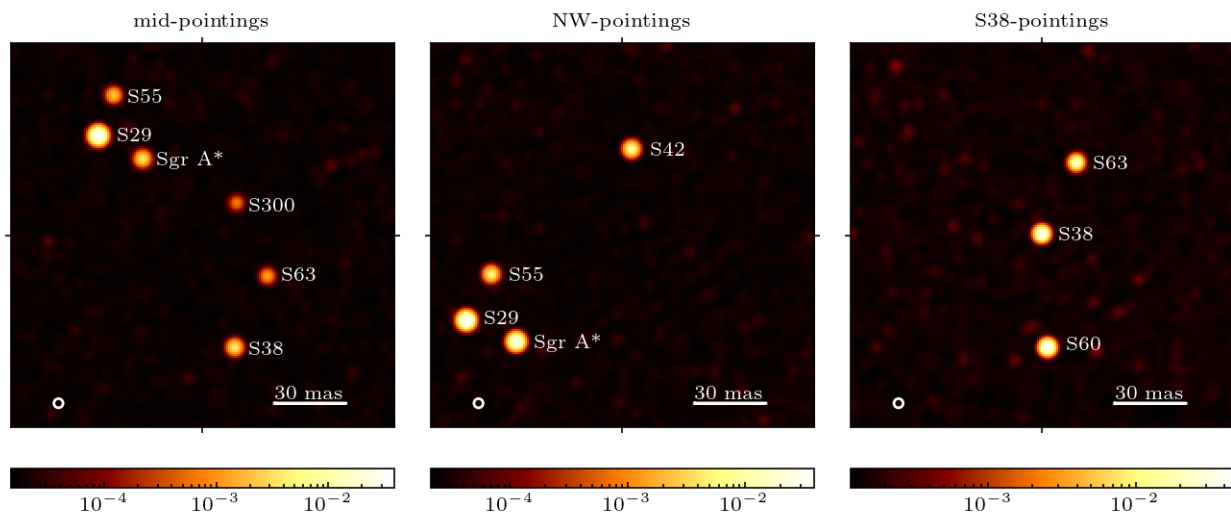
**Fig. 5.** Images obtained for the mosaicing data set, where north is up and east to the left. The flux is normalized by S29 in the mid and northwest pointings and by S38 in the rightmost panel. Images are computed as sample mean over all samples of a single $G^R$ run and have been convolved with a Gaussian of 1.6 mas standard deviation whose FWHM is indicated in the bottom left corner. The mid- and the S38-pointings extend the Sgr A*-centered images to the south west. S62 is not detectable in these images due to a combination of two effects. First, at this large distance from the image center, the already faint flux is further damped by the fiber profile; and secondly, the overall sensitivity of the images is reduced in comparison to the Sgr A* pointings because of the smaller number of exposures available.

**Table 5.** Orbital elements for six stars detected in the 2021 images.

| Star | $a$[mas] | $e$ | $i$ [°] | $\Omega$ [°] |
|------|----------|-----|---------|--------------|
| S2  | $124.95 \pm 0.04$ | $0.88441 \pm 0.00006$ | $134.70 \pm 0.03$ | $228.19 \pm 0.03$ |
| S29 | $397.50 \pm 1.56$ | $0.96927 \pm 0.00011$ | $144.37 \pm 0.07$ | $7.00 \pm 0.33$ |
| S38 | $142.54 \pm 0.04$ | $0.81451 \pm 0.00015$ | $166.65 \pm 0.40$ | $109.45 \pm 1.00$ |
| S42 | $411.42 \pm 7.14$ | $0.77385 \pm 0.00309$ | $39.57 \pm 0.19$ | $309.60 \pm 1.06$ |
| S55 | $104.40 \pm 0.05$ | $0.72669 \pm 0.00020$ | $158.52 \pm 0.22$ | $314.94 \pm 1.14$ |
| S60 | $489.21 \pm 18.55$ | $0.77733 \pm 0.00806$ | $126.60 \pm 0.15$ | $178.03 \pm 0.80$ |
| | $\omega$ [°] | $t_P$[yr] | $T$ [yr] | $m_K$ |
| S2  | $66.25 \pm 0.03$ | $2018.38 \pm 0.00$ | $16.046 \pm 0.001$ | $13.95$ |
| S29 | $205.79 \pm 0.33$ | $2021.41 \pm 0.00$ | $91.04 \pm 0.54$ | $16.6$ |
| S38 | $27.17 \pm 1.02$ | $2003.15 \pm 0.01$ | $19.55 \pm 0.01$ | $17.$ |
| S42 | $48.29 \pm 0.46$ | $2022.12 \pm 0.02$ | $95.9 \pm 2.5$ | $17.5$ |
| S55 | $322.78 \pm 1.13$ | $2009.44 \pm 0.01$ | $12.25 \pm 0.01$ | $17.5$ |
| S60 | $30.43 \pm 0.21$ | $2023.62 \pm 0.05$ | $124.3 \pm 7.1$ | $16.3$ |

**Notes.** The parameters listed are the semi major axis, eccentricity, inclination, position angle of the ascending node, longitude of periastron, epoch of periastron passage, orbital period, and K-band magnitude. The stars S62 and S300 are not listed here, as their movement observed so far is consistent with a linear motion. Images serve as starting point for high-accuracy astrometric fitting, and the orbits are determined from positions provided by the latter.

### 4.5. Sensitivity estimation

As a first step to estimate the sensitivity in our images, we performed a series of injection tests that are reported in Appendix F. They consider sources at two different locations and four magnitudes, between 19.7 and 22.7, which are inserted into the May 29 data set.

The ability of $G^R$ to recover the injected source depends on its position. In the first scenario, the new star is located close to S300, and we managed a high-significance detection only at 19.7th magnitude. In addition, we observed a larger scatter around S300, and more flux is placed in its side-lobes. In the second case, the source is injected at same distance but to the north east of Sgr A*. Here, we can reach significantly deeper and manage a robust detection even at 21.0th magnitude.

We can also use S300 itself to estimate the sensitivity in our images. As the star moves away from the field center, it is more strongly affected by fiber damping and appears fainter to the GRAVITY instrument. Since S300 is most robustly detected close to the field center, we use the images from March, May, and the mid pointing in July to estimate its magnitude from which we obtain $m_K(S300) \simeq 19.0 - 19.3$. The corresponding apparent magnitudes, that is, the magnitude corrected for fiber damping, are listed in Table 6 for all observing epochs.

Until June 2021, S300 is very robustly detected in our images which also matches the results of the mock tests above. In July, on the other hand, the apparent magnitude of S300 is already below 20, and we only managed a weak detection at low significance.

In addition to the heightened fiber damping, the large distance to the field center in combination with systematic effects can also affect the ability to detect S300 in July. In the context of our forward model, we can understand systematics as a mis-
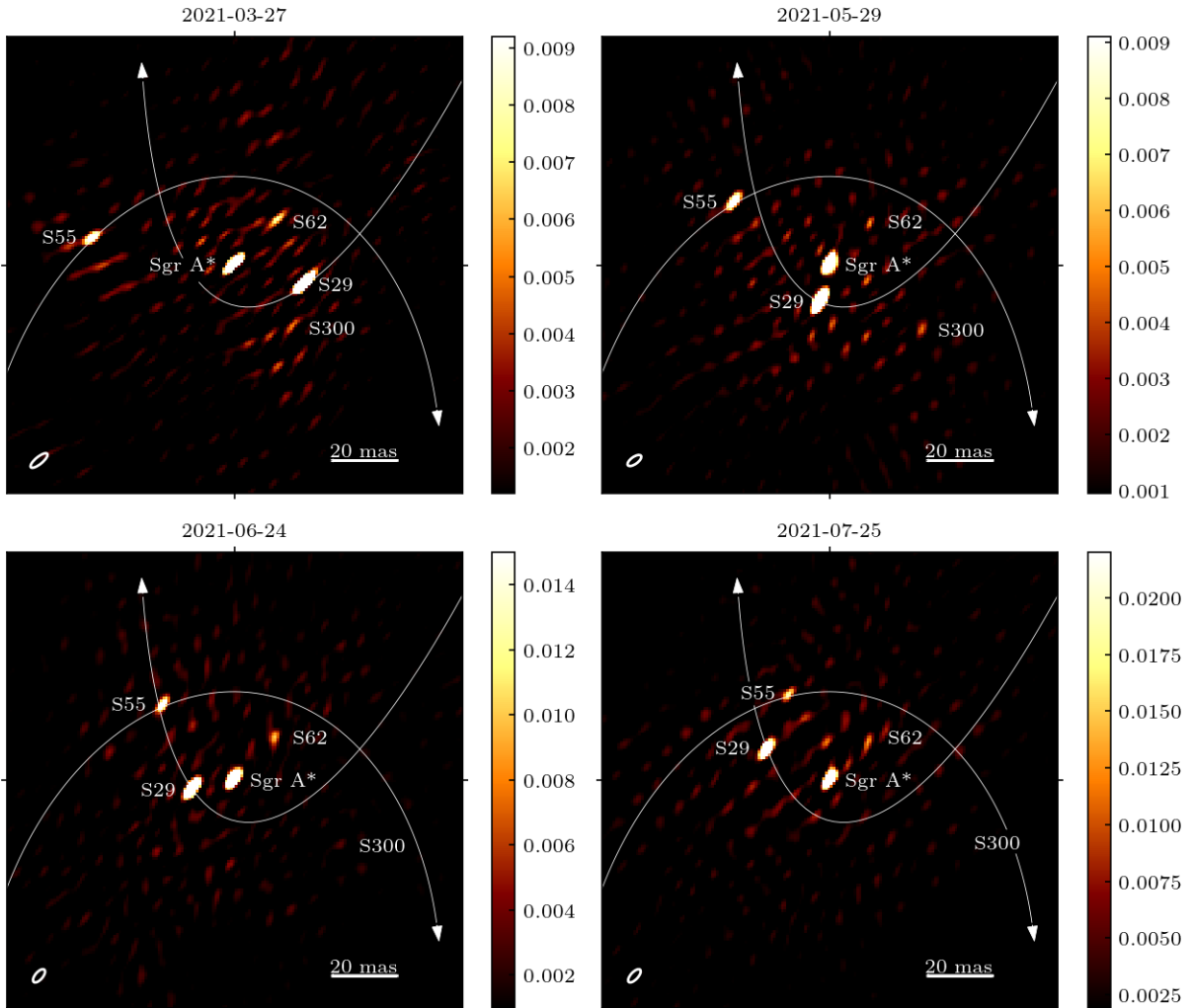
**Fig. 6.** Representative image for each observing epoch obtained with CLEAN (north is up, east to the left, and the flux is normalized to S2). We show the model convolved with the CLEAN beam on top of the residual images and indicate the FWHM of the CLEAN beam in the bottom left corner. To obtain the model, we cleaned Sgr A*, S29, S55, and S62; for the images from March to June, we also cleaned S300.

**Table 6.** Apparent magnitude of S300 in all four 2021 observing epochs.

| Date | min $m_{K,app}$ | max $m_{K,app}$ |
|---|---|---|
| March 2021 | 19.3 | 19.7 |
| May 2021 | 19.6 | 19.9 |
| June 2021 | 19.8 | 20.1 |
| July 2021 | 20.1 | 20.5 |

**Notes.** The apparent magnitude of a star accounts for corrections due to fiber damping if it is not located at the center of the FOV. The minimum and maximum values correspond to our bracketing estimates for the intrinsic S300 brightness.

match between the predicted signal of the source and its actual effect on the data. It is expected that the quality of our modeling decreases for sources further away from the field center. The possible reasons for this include bandwidth smearing and optical aberrations in the instrument. Both effects become stronger further off axis, and, consequently, their modeling is more sensitive to approximations and calibration data, such as the bandpass shape or the aberration maps. It is important to note in this context, that Bayesian analyses in high dimensions are particularly vulnerable to model mismatches and that alternative tools can

have a higher degree of robustness under such circumstances. We come back to this issue in Sect. 5.1, when we compare the new imaging algorithm to CLEAN.

At this point, we have encountered several factors beyond $(u, v)$-coverage and data quality that can considerably impact the sensitivity of our images, such as the brightness of Sgr A*, proximity to another faint source and the distance to the image center. Rather than giving a single estimate for the limiting magnitude, we therefore decided to highlight two bracketing values. Even under somewhat difficult circumstances – in June, S300 is already seen to be outside the FWHM of the fiber and in the first mock test, the injected source is very close to S300 – we are able to recover a source of at least $m_K \simeq 19.7$ magnitude when corrected for fiber damping. On the other hand, mock tests at position 2, and also the low noise levels reported in Appendix G, indicate that under favorable circumstances, we are able to push the sensitivity significantly beyond a magnitude of 20 with the newly developed $G^R$ imaging algorithm.

# 5. Discussion

## 5.1. Confirmation of S300 with CLEAN

So far, the standard for deep imaging of the GC with GRAVITY has been set by CLEAN (Gravity Collaboration et al. 2021a), and comparing both methods is important to judge the performance of our new $G^R$ algorithm. We therefore carried out an independent analysis of the 2021 data with CLEAN (using the AIPS implementation, Greisen 2003), which was already informed of the detection of S300 with $G^R$. In Fig. 6, we provide a representative image for each observing epoch and in Table 7, we list the corresponding data sets.

To successfully apply CLEAN to GRAVITY GC observations, a distinct procedure was developed and described in Gravity Collaboration et al. (2021a). Here, we adapt it to the changed field configuration in 2021. The major steps are laid out in the following.

Just as in 2019, we start by computing the coherent flux from the complex visibilities and the photometric flux observed at each telescope. To correct for flux variations (introduced e.g., by changes in air mass or the performance of the adaptive optics system), we interpolated the photometric flux of S2 across all frames collected for a night and normalized the coherent flux in the Sgr A*-centered pointings by this value.

We then cleaned on Sgr A* and S29 on an exposure-by-exposure basis. In particular, S29 is the brightest source in the field, but cleaning on Sgr A* is important due to its flux variations, which would otherwise render the data mutually inconsistent. After this has been resolved, we combine all Sgr A*- and S29-cleaned exposures from the night. In the resulting residual image, S55 is clearly visible and we jointly clean on the remains of Sgr A* and S29, as well as on S55. At this point, S62 usually becomes apparent. For three of the four nights shown in Fig. 6, it was the brightest residual, and it was only in the May imaging that it appeared as the second brightest.

The images obtained with CLEAN (cf. Fig. 6) agree very well with the $G^R$ results (cf. Fig. 1). Both methods recover an identical image structure that is dominated by the four established stars and matching source positions. Further, the CLEAN images confirm the detection of S300. For the March and May nights shown in Fig. 6, after S62 has been cleaned, a bright residual becomes visible whose location is consistent with the $G^R$ discovery. It corresponds to the second-brightest residual on March 27 and the third-brightest on May 29. Also on the June night, a residual is visible at the expected S300 position; however, in this case, it is not among the few brightest ones. Here, fiber damping considerably impacts the ability to detect S300 in the CLEAN images. Bandwidth smearing, which becomes more severe for sources further away from the image center and is not modeled by CLEAN, can also diminish the sensitivity to S300 for the later 2021 observing runs. Finally, the images in Fig. 6 were obtained by cleaning on S300.

The most powerful confirmation of S300 in the CLEAN images, however, is provided by the July mid pointings in Fig. 7. The image is obtained from a total of 16 frames collected over three consecutive nights from July 25 to July 27, 2021. After cleaning on Sgr A* in each individual exposure and combining all files, S38, S63, and S300 become apparent in the residual image, and we clean on the sources in this order. Subsequently, S29 can be recognized, which appears fainter in the mid pointings due to the larger fiber damping, and we also cleaned on it. Again, the CLEAN image shows excellent consistency with the $G^R$ result in Fig. 5.

**Table 7.** Summary of the data set from which the CLEAN images in Fig. 6 were obtained.

| Date | $N_{exp}$ | rms/S2 | corresponding $m_K$ |
|---|---|---|---|
| 2021-03-27 | 12 | $1.5 \times 10^{-3}$ | 21.2 |
| 2021-05-29 | 20 | $1.2 \times 10^{-3}$ | 21.4 |
| 2021-06-24 | 32 | $1.4 \times 10^{-3}$ | 21.2 |
| 2021-07-25 | 28 | $2.5 \times 10^{-3}$ | 20.6 |

**Notes.** Here, $N_{exp}$ gives the number of exposures used for imaging. In addition, we also provide the rms in the residual image normalized to S2 and translate the value to a K-band magnitude in the final column.

As in Gravity Collaboration et al. (2021a), we computed the root mean square (rms) value of the residual image in the central $74 \times 74$ mas for the Sgr A* centered images, see Table 7. The rms of the complex visibilities and the residual image are directly related by Parseval's theorem, such that the latter quantity measures how well the model is able to reproduce the data. The flux in the CLEAN images as well as the rms values are normalized to S2, a 14.1 magnitude star in K-band (Gravity Collaboration et al. 2017). Equivalent values for the $G^R$ results are estimated in Appendix G and are given in Table G.1.

The low brightness of S300 in the June CLEAN residual image, where $G^R$ still manages a robust detection, already illustrates the increased sensitivity of the latter method. Furthermore, for the May and June observing nights, where we can directly compare the results in Table 7 and Table G.1, $G^R$ improves the rms over the CLEAN result by $\Delta m_K \simeq 0.3 - 0.4$. At this point, we want to emphasize that the comparison of residual images is rather unfavorable for $G^R$, which reconstructs the image from closure phases and visibility amplitudes. Residual images are only computed retroactively and require some additional alignment steps which are described in Appendix G and can increase the noise level by themselves.

Apart from being better able to describe instrumental effects, $G^R$ has another advantage. The residual images are essential within CLEAN to search for new sources. They do not, however, contain any information about the error bars in the data domain. Even if the residual visibilities were completely consistent with the Gaussian noise expectation, the inhomogeneous sampling pattern of the Fourier plane would introduce structures in the residual images. It can be challenging to judge whether a bright spot in the residual image is consistent with noise expectations or corresponds to a faint source. In contrast, $G^R$ judges the match between model and data directly in the domain of the data and thus can compare the residuals directly with the error bars of the individual data points.

On the other hand, a clear advantage of CLEAN is its speed and the significantly smaller computational demands in comparison to $G^R$. A full image reconstruction starting from data conversion with CLEAN only requires a single CPU and takes one to one and a half hours, which are dominated by data selection and inspecting the results rather than computation time. The $G^R$-code runs for about ten times as long on 16 CPUs, albeit once it has been initiated, it requires no human intervention.

Finally, we note that in some situations, CLEAN appears to be more robust to shortcomings of the response model. For the mid-pointing images with $G^R$ in Sect. 4.4, we only used data from a single night. Combining multiple nights with $G^R$ works well for S2 and S38 pointings, but this significantly degrades the imaging solution of the mid pointings by introducing spurious sources, most likely due to the day-to-day movement of S29 and S55. While this further illustrates the heightened sensitivity of
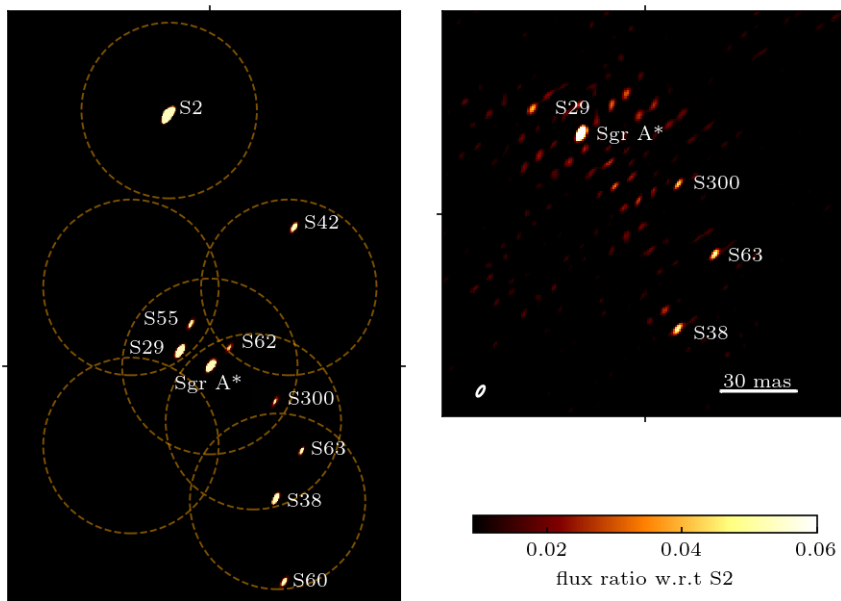
**Fig. 7.** Imaging results for the wide field data set obtained with CLEAN, where north is up, east to the left, and the flux is normalized to S2. In the left panel, we combine the CLEAN model from all pointings in the July 2021 mosaicing data set (see Table 2) and indicate the pointing direction by a circle with 50 mas radius. Apart from the expected stars, which are also summarized in Fig. 4, we do not detect any additional sources. In the right panel, we show the mid pointings from July 25, 26, and 27 (16 exposures in total) imaged with CLEAN. The image displays the model convolved with the CLEAN beam on top of the residuals and the beam size is indicated in the bottom left corner. To obtain the model, we have cleaned on Sgr A*, S38, S63, S300, and S29.

$G^R$, it also implies that additional work on the prior model will be required before one can combine exposures with fast moving sources from multiple nights. CLEAN, on the other hand, is able to image mid pointings from three consecutive nights jointly, without any artifacts from the movement of fainter stars affecting the structure of the brightest sources. In Fig. 7, we have Sgr A*, S300, S63, and S38 all correctly recovered, while the fast-moving S29, which is subdominant to the total flux, appears slightly smeared out at an average position. On the other hand, S55, which is even fainter and further off axis than S29, cannot be recovered.

### 5.2. Possible S300 positions in the Milky Way

The observed S300 positions from March to July (cf. Table 4 and Sect. 4.4) are fully consistent with a linear motion. Fitting for the angular velocity, we obtain:

$$v_{RA} = (-65.6 \pm 1.8) \text{ mas/yr},$$
$$v_{Dec} = (-0.5 \pm 2.7) \text{ mas/yr}.$$

Its high angular velocity makes it very unlikely that S300 is a background star. In the following, we discuss possible options for describing its nature, namely, as a star located in the CG and a foreground star in the galactic disk.

If the star is at the same distance as the GC, its projected velocity with respect to the GC would be $v \simeq 2.6 \times 10^3$ km/s. The $2\sigma$ limit on the acceleration of S300 is $-23$ mas/yr$^2$. This, in combination with requiring it to be gravitational bound to the black hole, limits the possible range of perpendicular coordinates to $50$ mas $\leq |z| \leq 137$ mas. Furthermore, for each possible value of $z$ there is a maximum velocity $v_z$ beyond which the star would become unbound. Sampling uniformly from the allowed values for $z$ and their corresponding velocities, we can investigate the possible orbits of S300.

We find that the median orbit of this distribution has a 171 mas semi-major axis, 0.6 eccentricity, and a 26 yr orbital period. It thus perfectly fits into the distribution of S-stars. Assuming a similarly accurate position determination as for the 2021 imaging, a single good observing night from 2022 would allow

to detect the acceleration of this median orbit. However, the distribution also contains more extreme solutions with larger semi-major axis and orbital period. For them, the apparently linear motion could continue throughout 2022. Even though, the continued observation of S300 over the coming year will allow to considerably constrain the family of allowed orbits.

Until then, in the absence of a significant acceleration detection, the possibility remains that S300 is a foreground star rather than being located in the GC. In this case, there would be little dust attenuation and S300 should also be detectable in the optical. We therefore checked the Gaia EDR3 catalog (Gaia Collaboration et al. 2021), which lists sources down to a G-band magnitude of 21, but we do not obtain a match. In a circle with 4 arc min diameter around Sgr A*, Gaia lists 285 stars fainter than $m_G > 17$. If we use this number at face value to estimate the number density of stars towards the GC, the probability for a star crossing our $100 \times 100$ mas image is small $\simeq 6 \times 10^{-5}$.

Further, if S300 was a disk star, there would be only a narrow range of possible distances from the sun consistent with our observations. While the high angular velocity implies that S300 must be close, the linear motion and the absence of a parallax also impose a minimum distance. Taken together, in addition to the low probability of a foreground star crossing the narrow GRAVITY FOV, its distance to the sun is limited to $O$ (kpc).

## 6. Conclusions

In this paper, we present GRAVITY-RESOLVE ($G^R$), a new imaging code, specifically tailored to GRAVITY observations of the Galactic Center (GC). The tool is based on a Bayesian interpretation of the imaging process and builds upon RESOLVE (Arras et al. 2021a; Arras et al. 2018), an imaging tool for radio interferometry developed in the framework of information field theory (Enßlin 2019). In this context, we implemented an instrument model which accounts for all relevant effects in GRAVITY and developed a prior that is specifically designed for the GC and can, for instance, accommodate the variability of Sgr A*. The posterior exploration was performed with Metric Gaussian Variational Inference (MGVI, Knollmüller & Enßlin 2019).

We then applied $G^R$ to GC observations in 2021. The resulting images reveal a complicated structure, composed of several point sources of different brightness with a time-variable central object. We note that while our prior model is specifically tailored to the GC and only considers point sources, there are also methods available to model an extended emission within a similar framework (Arras et al. 2020, 2021a).

The stars S29 and S55 (Gillessen et al. 2017) both pass their pericenters in 2021 and we detected them within a 50 mas radius from Sgr A*. Their position in the images is also the starting point for astrometric model fitting which allows to determine the sources' positions to a very high accuracy of $\sim$ 100 μas. We present a detailed study of the resulting GR orbits in a second publication (Gravity Collaboration 2021). Further, $G^R$ yields a very robust night-by-night detection of S62, a $m_K = 18.9$ star that slowly approaches Sgr A* and has already been found in GRAVITY GC images from 2019 obtained with CLEAN (Gravity Collaboration et al. 2021a). None of the sources S29, S55, S62, and S300 identified in the GRAVITY observations matches a 9.9 year orbital period star as reported in Peißker et al. (2020, 2021).

In addition to the known stars, a new source is apparent from the images which moves to the west at high angular speed $\simeq$ 66 mas/yr. It is detected at a high significance in the March, May, and June observations, but only dimly recognizable in the July Sgr A*-centered exposures, where it is located at the largest distance from the image center. In July, however, we performed some dedicated mid pointings in which the GRAVITY fibers are offset from Sgr A* and indeed recover the star at the expected location.

The new source neither corresponds to any of the known S-stars (Gillessen et al. 2017) nor is it present in the Gaia catalog (Gaia Collaboration et al. 2021), and we refer to it as S300. From the flux ratio with known stars in the field, particularly S29 and S38, we estimate that $m_K(S300) = 19.0 - 19.3$. If fiber damping is taken into account, this source is at the detection limit estimated in Gravity Collaboration et al. (2021a) in the March images and becomes dimmer during the rest of the year. The detection clearly demonstrates that $G^R$ can deliver significantly deeper images. With the knowledge of the S300 position, we also searched for it in 2021 CLEAN images and, indeed, we can identify a bright residual at the correct location in March, May, June, and the mid pointings.

The mid pointings are part of a larger mosaicing data set (cf. Table 2), obtained to scan a wider field around Sgr A*. In the corresponding images (cf. Fig 5), we can detect S38, S42, S60, and S63 in addition to the aforementioned sources. We use the opportunity to give an update on the orbital elements of all stars in Table 5. These are based on astrometric fits to the data for which the imaging serves as a starting point.

Struggles to detect S300 in the July Sgr A*-centered frames are partly due to fiber damping, but also the fact that our instrumental model becomes more sensitive to approximations; in addition, uncertainties in the calibration the farther away from the image center a source is located might play a role. To further assess the sensitivity of the new code, we performed a series of injection tests that we compared to the apparent magnitude of S300. As the images from May 22 demonstrate, a flare significantly reduces our sensitivity to faint sources. Under moderately difficult circumstances – if the source is injected close to S300 or in June, where the off-axis separation already exceeds the fiber FWHM – we are still able to robustly recover a source with an apparent magnitude of at least 19.7. Finally, the injection tests show that, under good circumstances, we are able to push the sensitivity significantly below a magnitude of 20 and we are even able to retrieve a injected star of a magnitude of 21.0.

At present, a limiting factor to $G^R$ is that we have to scale the error bars by hand and to adjust that scaling by trial-and-error, making the application cumbersome. This can be improved in the future by implementing an automated error inference, such as in Arras et al. (2019b). With this method, we can also introduce more elaborate covariance matrices for the likelihood. Indeed, we find that the correlation structure of our residuals is similar to the one reported in Kammerer et al. (2020). This analysis, which was carried out in the context of exo-planet observations with GRAVITY, demonstrated that an improved correlation model increases the achievable contrast. We see a similar potential in the imaging context, which would allow for an improvement in the convergence of the inference and further enhance the sensitivity beyond the capabilities that we have demonstrated with this publication.

# References

Arras, P., Baltac, M., Ensslin, T. A., et al. 2019a, NIFTy5: Numerical Information Field Theory v5

Arras, P., Bester, H. L., Perley, R. A., et al. 2021a, A&A, 646, A84

Arras, P., Frank, P., Haim, P., et al. 2020, arXiv e-prints, arXiv:2002.05218

Arras, P., Frank, P., Leike, R., Westermann, R., & Enßlin, T. A. 2019b, A&A, 627, A134

Arras, P., Knollrnüller, J., Junklewitz, H., & Enßlin, T. A. 2018, in 2018 26th European Signal Processing Conference (EUSIPCO), 2683–2687

Arras, P., Reinecke, M., Westermann, R., & Enßin, T. A. 2021b, A&A, 646, A58

Baron, F., Monnier, J. D., & Kloppenborg, B. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7734, Optical and Infrared Interferometry II, ed. W. C. Danchi, F. Delplancke, & J. K. Rajagopal, 77342I

Bhatnagar, S., Cornwell, T. J., Golap, K., & Uson, J. M. 2008, A&A, 487, 419

Blackburn, L., Pesce, D. W., Johnson, M. D., et al. 2020, ApJ, 894, 31

Briggs, D. S., Schwab, F. R., & Sramek, R. A. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 180, Synthesis Imaging in Radio Astronomy II, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley, 127

Do, T., Ghez, A. M., Morris, M. R., et al. 2009, ApJ, 691, 1021

Do, T., Hees, A., Ghez, A., et al 2019, Science, 365, 664

Do, T., Lu, J. R., Ghez, A. M., et al. 2013, ApJ, 764, 154

Dodds-Eden, K., Gillessen, S., Fritz, T. K., et al. 2011, ApJ, 728, 37

Dutt, A. & Rokhlin, V. 1993, SIAM Journal on Scientific Computing, 14, 1368

Eckart, A., Schödel, R., García-Marín, M., et al. 2008, A&A, 492, 337

Eisenhauer, F., Genzel, R., Alexander, T., et al. 2005, ApJ, 628, 246

Enßlin, T. A. 2019, Annalen der Physik, 531, 1970017

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1

Gallego-Cano, E., Schödel, R., Dong, H., et al. 2018, A&A, 609, A26

Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, Reviews of Modern Physics, 82, 3121

Genzel, R., Schödel, R., Ott, T., et al. 2003a, Nature, 425, 934

Genzel, R., Schödel, R., Ott, T., et al. 2003b, ApJ, 594, 812

Ghez, A. M., Wright, S. A., Matthews, K., et al. 2004, ApJ, 601, L159

Gillessen, S., Eisenhauer, F., Quataert, E., et al. 2006, ApJ, 640, L163

Gillessen, S., Eisenhauer, F., Trippe, S., et al. 2009, ApJ, 692, 1075

Gillessen, S., Plewa, P. M., Eisenhauer, F., et al. 2017, ApJ, 837, 30

Gravity Collaboration. 2021, A&A

Gravity Collaboration, Abuter, R., Accardo, M., et al. 2017, A&A, 602, A94

Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018a, A&A, 615, L15

Gravity Collaboration, Abuter, R., Amorim, A., et al. 2020a, A&A, 638, A2
Gravity Collaboration, Abuter, R., Amorim, A., et al. 2020b, A&A, 636, L5
Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018b, A&A, 618, L10
Gravity Collaboration, Abuter, R., Amorim, A., et al. 2021a, A&A, 645, A127
Gravity Collaboration, Abuter, R., Amorim, A., et al. 2021b, A&A, 647, A59
Gravity Collaboration, Bauböck, M., Dexter, J., et al. 2020c, A&A, 635, A143
Gravity Collaboration, Jiménez-Rosales, A., Dexter, J., et al. 2020d, A&A, 643, A56
Greisen, E. W. 2003, AIPS, the VLA, and the VLBA, ed. A. Heck, Vol. 285, 109
Guyon, O. 2002, A&A, 387, 366
Högbom, J. A. 1974, A&AS, 15, 417
Kammerer, J., Mérand, A., Ireland, M. J., & Lacour, S. 2020, A&A, 644, A110
Knollmüller, J. & Enßlin, T. A. 2019, arXiv e-prints, arXiv:1901.11033
Merritt, D., Alexander, T., Mikkola, S., & Will, C. M. 2010, Phys. Rev. D, 81, 062002
Meyer, L., Ghez, A. M., Schödel, R., et al. 2012, Science, 338, 84
Peißker, F., Eckart, A., & Ali, B. 2021, ApJ, 918, 25
Peißker, F., Eckart, A., & Parsa, M. 2020, ApJ, 889, 61
Perrin, G. & Woillez, J. 2019, A&A, 625, A48
Rogers, A. E. E., Hinteregger, H. F., Whitney, A. R., et al. 1974, ApJ, 193, 293
Shahzamanian, B., Eckart, A., Valencia-S., M., et al. 2015, A&A, 576, A20
Shaklan, S. & Roddier, F. 1988, Appl. Opt., 27, 2334
Smirnov, O. M. 2011, A&A, 527, A106
Thiébaut, E. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7013, Optical and Infrared Interferometry, ed. M. Schöller, W. C. Danchi, & F. Delplancke, 70131I
Trippe, S., Paumard, T., Ott, T., et al. 2007, MNRAS, 375, 764
van Cittert, P. H. 1934, Physica, 1, 201
Waisberg, I., Dexter, J., Gillessen, S., et al. 2018, MNRAS, 476, 3600
Witzel, G., Martinez, G., Hora, J., et al. 2018, ApJ, 863, 15
Zamaninasab, M., Eckart, A., Witzel, G., et al. 2010, A&A, 510, A3
Zernike, F. 1938, Physica, 5, 785
Zhang, F. & Iorio, L. 2017, ApJ, 834, 198
Zucker, S., Alexander, T., Gillessen, S., Eisenhauer, F., & Genzel, R. 2006, ApJ, 639, L21

[1] Max Planck Institute for extraterrestrial Physics, Giessenbach-straße 1, 85748 Garching, Germany
[2] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195 Meudon, France
[3] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[4] 1st Institute of Physics, University of Cologne, Zülpicher Straße 77, 50937 Cologne, Germany
[5] Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[6] Universidade de Lisboa - Faculdade de Ciências, Campo Grande, 1749-016 Lisboa, Portugal
[7] Faculdade de Engenharia, Universidade do Porto, rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[8] European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching, Germany
[9] European Southern Observatory, Casilla 19001, Santiago 19, Chile
[10] Sterrewacht Leiden, Leiden University, Postbus 9513, 2300 RA Leiden, The Netherlands
[11] Departments of Physics and Astronomy, Le Conte Hall, University of California, Berkeley, CA 94720, USA
[12] CENTRA - Centro de Astrofísica e Gravitação, IST, Universidade de Lisboa, 1049-001 Lisboa, Portugal
[13] Department of Astrophysical & Planetary Sciences, JILA, Duane Physics Bldg., 2000 Colorado Ave, University of Colorado, Boulder, CO 80309, USA
[14] Department of Particle Physics & Astrophysics, Weizmann Institute of Science, Rehovot 76100, Israel
[15] Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
[16] Department of Physics, Technical University Munich, James-Franck-Straße 1, 85748 Garching, Germany
[17] Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[18] Department of Physics, University of Illinois, 1110 West Green Street, Urbana, IL 61801, USA
[19] Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
[20] CERN, 1 Esplanade des Particules, Genève 23, CH-1211, Switzerland
[21] Department of Astrophysics, IMAPP, Radboud University, 6500 GL Nijmegen, The Netherlands

## Appendix A: Efficient response implementation

In the presence of static instrumental optical aberrations (Gravity Collaboration et al. 2021b) and bandwidth smearing, the simplified response function in Eq. (3) is generalized to:

$$v_{ij}(\boldsymbol{u}) = \frac{\int d\lambda \, P_{\lambda_0}(\lambda) \int d\boldsymbol{s} \, \Pi_i(\boldsymbol{s}) \, \Pi_j^*(\boldsymbol{s}) \, I(\boldsymbol{s}, \lambda) \, e^{-2\pi i \boldsymbol{b} \cdot \boldsymbol{s}/\lambda}}{\prod_{x=i,j} \sqrt{\int d\lambda \, P_{\lambda_0}(\lambda) \int d\boldsymbol{s} \, |\Pi_x(\boldsymbol{s})|^2 \, I(\boldsymbol{s}, \lambda)}}, \quad \text{(A.1)}$$

where $i$ and $j$ label the two telescopes forming the baseline, $\boldsymbol{b}$, and the Fourier coordinate is evaluated at the central wavelength $\lambda_0$, that is, $\boldsymbol{u} = \boldsymbol{b}/\lambda_0$. The spectral bandpass is normalized such that $\int d\lambda \, P_{\lambda_0}(\lambda) = 1$. Finally, $\Pi_{i/j}$ summarizes the phase maps, $\phi_{i/j}$, and amplitude maps, $A_{i/j}$, of each telescope and is given by

$$\Pi_i(\boldsymbol{s}) = A_i(\boldsymbol{s} + \boldsymbol{\delta}_i) \, e^{i\phi_i(\boldsymbol{s}+\boldsymbol{\delta}_i)}. \quad \text{(A.2)}$$

Since the fiber center is not necessarily aligned perfectly with the image center, a small offset $\boldsymbol{\delta}_i$ can arise between the zero-coordinate of the phase and amplitude maps and the sky intensity distribution.

The position integral in the numerator of Eq. (A.1) can be computed very efficiently with the fast fourier transform (FFT) algorithm. However, the grid of rectangular frequencies does not necessarily align with the $(u, v)$-coordinates of the measurements. A common technique in radio and optical interferometry, known as gridding, is to interpolate the complex visibilities between the regular grid points of the FFT (Dutt & Rokhlin 1993; Briggs et al. 1999).

To implement the phase and amplitude maps numerically, we produced a complex screen for each telescope, following the procedure in Gravity Collaboration et al. (2021b). The maps already account for the instrument's aperture and are smoothed with a 10 mas Gaussian kernel that models residual tip-tilt jitter from the AO system (Perrin & Woillez 2019). The resulting complex screen has the same dimensions as the actual image, and both are multiplied pixel-wise before FFT and gridding are performed[2]. Because the aberrations differ between telescopes, we need to perform the computation separately on each baseline.

Taking into account the temporal and spectral variability naively would lead to a slightly different image for each exposure and each spectral channel and further curtail any computational gain achievable by gridding. To avoid this, we write the overall sky brightness distribution as

$$I(\boldsymbol{s}, \lambda, t, p) = f_{\nu_1}(\lambda) \, I_{\text{SgrA}}(t, p) \, \delta\left(\boldsymbol{s} - \boldsymbol{s}_{\text{SgrA}}\right)$$
$$+ f_{\nu_2}(\lambda) \left[ I_{\text{Img}}(\boldsymbol{s}) + \sum_{j=1}^{N_{\text{PS}}} I_j \, \delta\left(\boldsymbol{s} - \boldsymbol{s}_j\right) \right], \quad \text{(A.3)}$$

where the first line corresponds to Sgr A*, a variable point source with spectral index $\nu_1$ and the second line represents the image of faint sources plus $N_{\text{PS}}$ additional point sources whose spectral index is $\nu_2$. The spectral distributions are given by

$$f_\nu(\lambda) = \left(\frac{\lambda}{2.2\,\mu\text{m}}\right)^{-\nu-1}, \quad \text{(A.4)}$$

---

[2] We use the ducc0 library for FFT and gridding (Arras et al. 2021b) which provides C++17 code with a comprehensive Python interface, cf. https://gitlab.mpcdf.mpg.de/mtr/ducc

where we choose the reference wavelength at the center of the K-band observed by GRAVITY. In this context, the spectral indices correspond to the observed spectrum, namely, the intrinsic spectrum of the sources altered by interstellar absorption and reddening, the Earth's atmosphere, and instrumental transmission. However, the complex visibilities in Eq. (A.1) are primarily sensitive to the difference $\nu_1 - \nu_2$, rather than to the individual spectral indices due to the normalization by the total flux.

We then approximate the bandpass integration by an average over $N_\lambda$ points, distributed linearly over each channel such that the numerator of Eq. (A.1) becomes

$$\frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} P_{\lambda_0}(\lambda_i) \left\{ f_{\nu_1}(\lambda_i) \, I_{\text{SgrA}}(t, p) \, e^{-2\pi i \boldsymbol{b} \cdot \boldsymbol{s}_{\text{SgrA}}/\lambda_i} + \right.$$
$$\left. f_{\nu_2} \left[ \sum_{j=1}^{N_{\text{PS}}} I_j \, e^{-2\pi i \boldsymbol{b} \cdot \boldsymbol{s}_j/\lambda_i} + \int d\boldsymbol{s} \, I_{\text{Img}}(\boldsymbol{s}) \, e^{-2\pi i \boldsymbol{b} \cdot \boldsymbol{s}/\lambda_i} \right] \right\}. \quad \text{(A.5)}$$

Because we implemented the last integral as an FFT and a subsequent gridding operation, evaluating it for a fine-gridded set of $\lambda$ values is computationally not very expensive as long as the image inside the Fourier transform is mono-chromatic. The computational steps to evaluate Eq. (A.5) are summarized in the following.

1. We use a FFT followed by gridding to compute

$$\int d\boldsymbol{s} \, A_i(\boldsymbol{s}) \, A_j(\boldsymbol{s}) \, I_{\text{Img}}(\boldsymbol{s}) \, e^{-2\pi i \boldsymbol{b} \cdot \boldsymbol{s}/\lambda + i\phi_i(\boldsymbol{s}) - i\phi_j(\boldsymbol{s})} \quad \text{(A.6)}$$

simultaneously for all exposures, all spectral channels, and all fine-grained $\lambda$ values within a channel.
2. For all additional point sources, we obtain $e^{-2\pi i \boldsymbol{b} \cdot \boldsymbol{s}/\lambda}$ at each exposure, channel and sub-resolution in $\lambda$ and multiply it by the phase- and amplitude-maps at the position of the source. Hereby, we approximate the aberration maps at the sources' actual positions by the aberration maps at the center of their position prior. Since the maps are very flat on the scale of the position uncertainty, the error introduced by this approximation is small, however it significantly simplifies the computation.
3. We multiply the expression for the time-dependent point source by an exposure- and polarization-dependent brightness $I_{\text{SgrA}}(t, p)$ and all further point sources by their constant brightness $I_j$ which is independent of the spectral channel.
4. We then account for the spectral distribution by multiplying each source at each $\lambda$-coordinate with the appropriate value from Eq. (A.4). This multiplication is the same for all exposures.
5. Finally, we add up all components, multiply by the spectral bandpass and perform the wavelength integration or bandwidth smearing as an average over all fine-grained $\lambda$-steps within one channel.
6. The procedure is repeated independently for all six GRAVITY baselines due to the changing aberration maps between baselines.

This implementation is very flexible in that more complicated spectra than the power-law of Eq. (A.4) can easily be accommodated. It can also readily be generalized to accomodate additional spectral components and different bandpass shapes.

We use calibration measurements of the GRAVITY spectral transmission to investigate the effect of bandwidth smearing and to compare the full measured bandpass to a top hat approximation. We find that the top-hat affects the computed visibilities on

a level much smaller than the typical GRAVITY error bars and, therefore, we chose the approximation for this work. To determine the required sub-sampling steps per spectral channel, we iteratively increase $N_\lambda$ until the visibilities have converged to a stable value. This is the case by a good margin at $N_\lambda = 100$.

To implement the denominator term of Eq. (A.1), we note that the spectral and positional integration can be decoupled under the parametrization in Eq. (A.3). Furthermore, assuming a top-hat bandpass with center $\lambda_0$ and width $2\delta\lambda$:

$$P(\lambda) = \text{rect}\left(\frac{\lambda - \lambda_0}{2\delta\lambda}\right) = \begin{cases} 1 & \text{if } |\lambda - \lambda_0| < 0 \\ 1/2 & \text{if } |\lambda - \lambda_0| = 0 \\ 0 & \text{otherwise,} \end{cases} \quad (A.7)$$

the wavelength integral is evaluated analytically to:

$$\int d\lambda \, P(\lambda) \, f_\nu(\lambda) = \frac{2.2\,\mu\text{m}}{\nu}\left[\left(\frac{\lambda_0 - \delta\lambda}{2.2\,\mu\text{m}}\right)^{-\nu} - \left(\frac{\lambda_0 + \delta\lambda}{2.2\,\mu\text{m}}\right)^{-\nu}\right]. \quad (A.8)$$

The integral in the spatial directions, on the other hand, can be computed readily as a sum over all image pixels multiplied by the squared amplitude maps. To this, we add the intensity of potential extra point sources multiplied by the fiber damping. For the variable component, the spatial integral is simply given by the product between light curve and fiber damping. Either contribution is then multiplied by Eq. (A.8), then their summation gives the denominator of Eq. (A.1).

Once we have evaluated the deterministic part of the response function, we multiply it by the time- and baseline dependent amplitude calibration factor, introduced in Sect. 2.3.

With this setup, one full response evaluation, including all exposures, baselines, and channels of a typical observing night with 21 exposures takes 0.4 s on a single laptop CPU (Intel i7 1.90 GHz).

## Appendix B: Formal description of the inference scheme

Given the sky model from Sect. 2.1 (cf. also Eq. A.3) and the self-calibration approach for the visibility amplitudes described in Sect. 2.3, the full set of model parameters to be inferred is

$$\theta = \big\{I_{\text{Img}}(s),\, I_{\text{SgrA}}(t, p),\, s_{\text{SgrA}},\, I_1, ... I_{\text{NP}},\, s_1, ...$$
$$...s_{\text{NP}},\, \nu_1,\, \nu_2,\, C(t, b)\big\}. \quad (B.1)$$

For a clearer notation, we have written the spatial ($s$), temporal ($t$), baseline ($b$) and the polarization ($p$) dependence of these quantities as functional arguments, but we use them, in particular, to label the discretized coordinates. The degrees of freedom inherent to each of these parameters are counted in Eq. (5). The prior model for the image of faint sources is provided in Eq. (2), and for the remaining parameters we have:

$$\ln I_{\text{SgrA}}(t, p) \hookleftarrow \prod_{j=1}^{N_{\text{exp}}} \prod_{p=1,2} \mathcal{G}\left[\ln I_{\text{SgrA}}(t_j, p)\,\Big|\,0, 1\right], \quad (B.2)$$

$$s_{\text{SgrA}} \hookleftarrow \mathcal{G}\left(s_{\text{SgrA}}\,\big|\,\mu_{\text{SgrA}},\, \sigma_{\text{SgrA}}\mathbb{1}\right), \quad (B.3)$$

$$\ln I_i \hookleftarrow \mathcal{G}(\ln I_i|\, 0, 1) \quad \text{for} \quad i = 1, ... N_{\text{PS}}, \quad (B.4)$$

$$s_i \hookleftarrow \mathcal{G}(s_i|\mu_i, \sigma_i\mathbb{1}) \quad \text{for} \quad i = 1, ... N_{\text{PS}}, \quad (B.5)$$

$$\nu_i \hookleftarrow \mathcal{G}(\nu_i|\mu_{\nu_i}, \sigma_{\nu_i}) \quad \text{for} \quad i = 1, 2, \quad (B.6)$$

$$C_i(t, b) \hookleftarrow \prod_{b=1}^{6} \prod_{j=1}^{N_{\text{exp}}} \mathcal{G}\left[C_i(t_j)\,\Big|\,1,\, \sigma_C\right], \quad (B.7)$$

where the "$\hookleftarrow$" indicates the probability density function from which a variable is drawn a priori and $N_{\text{exp}}$ counts the exposures in the data set. MGVI works with standardized coordinates, that is, all the model parameters are mapped to a set of auxiliary parameters $\xi$, which follow a standard normal distribution. This mapping is obtained by a two-step process,

$$\theta = \left[\mathcal{F}_{\mathcal{P}(\theta)}^{-1} \circ \mathcal{F}_{\mathcal{G}(\xi|0,\mathbb{1})}\right](\xi), \quad (B.8)$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the cumulative probability density function and its inverse, respectively. The first step translates a standard normal distributed to an uniformly distributed variable. In the second step, this uniform distribution is transformed to the prior specified in Eq. (2) and Eqs. (B.2) to (B.7).

For brevity, we introduce an operator notation in which $\mathcal{S}$ is the signal operator and corresponds to the application of Eq. (B.8). The response operator, $\mathcal{R}$, computes complex visibilities for any realization of $\theta$ as specified by Eq. (A.1). Finally, the complex visibilities are transformed to visibility amplitudes and closure phases by the $\mathcal{A}$ and $\mathcal{C}$ operator, respectively.

In this notation, the negative logarithm of the likelihood discussed in Sect. 2.4 is expressed as:

$$-\ln \mathcal{P}(d|\xi) \stackrel{\frown}{=} \frac{1}{2}\left[e^{i(C\circ\mathcal{R}\circ\mathcal{S})(\xi)} - e^{iC(d)}\right]N_\Phi^{-1}\left[e^{i(C\circ\mathcal{R}\circ\mathcal{S})(\xi)} - e^{iC(d)}\right]^\dagger$$
$$+ \frac{1}{2}\left[(\mathcal{A}\circ\mathcal{R}\circ\mathcal{S})(\xi) - \mathcal{A}(d)\right]N_\rho^{-1}\left[(\mathcal{A}\circ\mathcal{R}\circ\mathcal{S})(\xi) - \mathcal{A}(d)\right]^\dagger, \quad (B.9)$$

where "$\stackrel{\frown}{=}$" denotes equality up to addition of a constant and $N_\Phi^{-1}$ and $N_\rho^{-1}$ are the diagonal covariance matrices of the closure phases and visibility amplitudes (see Sect. 2.4). In the first line, we approximate the closure phases by their position on the unit circle, which mitigates the problem of phase wraps. This represents an approximation to the full likelihood, which becomes precise in the limit of small differences and down-weights outliers. The corresponding negative log-prior, on the other hand, is now simply given by

$$-\ln \mathcal{P}(\xi) \stackrel{\frown}{=} \frac{1}{2}\xi\mathbb{1}\xi^\dagger. \quad (B.10)$$

In the inference, MGVI approximates the full posterior by a multivariate Gaussian distribution with covariance $\Xi \simeq I(\xi)^{-1}$, where $I(\xi)$ is a generalization of the Fisher metric defined as

$$I(\xi) = \mathbb{1} + \left[\frac{\partial\, e^{i(C\circ\mathcal{R}\circ\mathcal{S})(\xi)}}{\partial\xi}\right]N_\Phi^{-1}\left[\frac{\partial\, e^{i(C\circ\mathcal{R}\circ\mathcal{S})(\xi)}}{\partial\xi}\right]^\dagger$$
$$+ \left[\frac{\partial\,(\mathcal{A}\circ\mathcal{R}\circ\mathcal{S})(\xi)}{\partial\xi}\right]N_\rho^{-1}\left[\frac{\partial\,(\mathcal{A}\circ\mathcal{R}\circ\mathcal{S})(\xi)}{\partial\xi}\right]^\dagger. \quad (B.11)$$

Here, the first term originates from the prior and the latter two from the likelihood. The NIFTy package (Arras et al. 2019a), which we have used to implement the inference, facilitates autodifferentiation, such that only the derivatives of individual operators need to be implemented and an operator sequence can be differentiated automatically by applying the chain rule.

## Appendix C: Hyper-parameters of the inference

The performance of MGVI depends on a series of hyper-parameters which need to be set by the user. Among them is the number of MGVI iterations and the number of samples and minimization steps at each iteration which we jointly refer to as

**Table C.1.** Iteration scheme for the posterior exploration with MGVI.

| iteration step | # minimization steps | # samples |
|:---:|:---:|:---:|
| 1 - 3 | 4 | 3 |
| 4 & 5 | 5 | 5 |
| 7 & 8 | 8 | 6 |
| 9 - 12 | 10 | 8 |
| 13 - 20 | 15 | 12 |
| 21 - 25 | 20 | 15 |
| 26 - 30 | 30 | 20 |

the iteration scheme. In addition, it is possible to choose between different minimizers and iteration controllers for the minimization of the KL and for drawing samples at a given position.

We determined these hyper-parameters in a two step procedure and first consider a mock data set, for which we took a representative $(u, v)$-coverage from the 2019 observing series. We then specified a model with a variable Sgr A* at the center, two faint sources in its vicinity and a brighter S2-like source to the north east. The model provides a prediction for closure phases and visibility amplitudes, to which we added a random Gaussian noise realization.

In the analysis of the mock data set, we varied all the aforementioned hyper-parameters with the aim to optimally recover the ground truth. We find that an iteration controller that checks the relative energy change between steps works best for drawing samples and minimizing the KL, and for the latter task we picked a Newton Conjugate Gradient minimizer. None of our results show significant changes after 30 MGVI iterations.

Arras et al. (2020) noted that tempering during the early MGVI iterations could improve their results and to this end, they iterated between minimizing the closure and amplitude likelihoods separately. We do not find a similar improvement in our reconstructions and thus consider the full likelihood at all MGVI steps.

The performance on the actual data can differ from the mock-imaging tests, for example, if not all noise correlations are modeled or if some residual systematic effects remain unaccounted for by the response function. Therefore, in a second step, we revised our iteration scheme now considering actual observations. This resulted in an overall decrease in the number of minimization steps and an increase in the number of samples at each MGVI iteration. Accordingly, MGVI reaches the minimum more slowly and overfitting is impeded. We can thus view this revision as the adoption of a less aggressive and more conservative approach.

We summarize the minimization scheme in Table C.1. The consistency of our images across multiple nights and observing periods (cf. Sect. 4.3) and injection tests (cf. Sect. 4.5) verify that our result is not an artifact of the minimization routine.

# Appendix D: Identification of stars

The GRAVITY images illustrate how S29 flies through pericenter in 2021 and here we comment on the crossidentification of the star with the earlier, AO-based NACO data. For that purpose, we differentiate between S29$_{GRAVITY}$ (i.e., the star labeled S29 throughout this publication and observed with GRAVITY) and S29$_{NACO}$, the star label S29 in Gillessen et al. (2009).

## Appendix D.1: Identifying S29$_{GRAVITY}$ with S29$_{NACO}$

In Fig. D.1, we show a compilation of deconvolved NACO images from 2002 to 2019, which allow us to trace S29$_{NACO}$ through the years. In 2019, the star was observed with NACO and GRAVITY, and we obtain matching positions for S29$_{NACO}$ and S29$_{GRAVITY}$.

Using only the 2021 GRAVITY data for S29$_{GRAVITY}$, we can fit an orbit and extrapolate back to 2019, yielding a position consistent with the 2019 astrometric points of GRAVITY for S29$_{GRAVITY}$. Furthermore, the orbit from the full interfermetric data set, that is, GRAVITY astrometric positions from 2021 to 2019, reliably predicts the S29$_{NACO}$ position in earlier AO epochs. In particular, it agrees with the 2007 reference map in Gillessen et al. (2009), as demonstrated in Fig. D.2.

We conclude that our identification of S29$_{GRAVITY}$ with S29$_{NACO}$ is robust and can be done both forwards from the AO epochs to GRAVITY as well as backwards from the GRAVITY data to the NACO images obtained $15 - 20$ years earlier.

## Appendix D.2: Excluding other identifications for S29$_{GRAVITY}$

Conserved dynamical quantities can be used to tag stars, since the conserved quantity needs to apply at each point on the orbit for a given star. Here, we use the $z$-component of the angular momentum vector, $h_z = xv_y - yv_x$, which can be derived from astrometric data alone. For S29$_{GRAVITY}$, we can determine that value for four occasions in 2021: during each of the four observing campaigns, we were not only able to determine the positions, but also the proper motions. This yields $h_z = (-1.09\pm0.03) \times 10^{20}$ m$^2$/s, where the value is the mean over the four epochs and the error the corresponding standard deviation. The two 2019 points for S29$_{GRAVITY}$ yield $h_z \approx -1.1 \times 10^{20}$ m$^2$/s, consistent with the 2021 value.

This excludes, for example, that S29$_{GRAVITY}$ is identical to the source candidate from Peißker et al. (2021), which we henceforth refer to as S62$_{Peissker}$. The latter star has a significantly smaller z-component for the angular momentum vector, $h_z \approx +1.4 \times 10^{19}$ m$^2$/s. Indeed, a star with such a high value of $h_z$ as S29$_{GRAVITY}$ would need to show an (tangential) on-sky motion of $\approx 880$ km/s, for a projected radius of 0.1 mas, at appocenter.

We also inspected the two 2019 data sets obtained on S29$_{GRAVITY}$. Using the same imaging technique as in Sect. 5.1 and Gravity Collaboration et al. (2021a), we find that there is only one dominant source in the field. Furthermore, the residual level is fainter than $m_K = 19$. The structure of these residuals can be attributed to S2, which is significantly brighter than S29$_{GRAVITY}$ but located further away from the field center and thus appears significantly damped by the Gaussian acceptance profile of the GRAVITY fibers.

## Appendix D.3: Looking for the counterpart of S62$_{Peissker}$

There are two sets of GRAVITY pointings which contain the position predicted by the S62$_{Peissker}$ orbit in Peißker et al. (2020). First, we would have expected it to show up in the 2019 GRAVITY pointings to S29$_{GRAVITY}$. Second, we pointed to the northwest of Sgr A* this year as part of the mosaicing data set (see Table 2 and Fig. 4). In the former, we detect S29 at the expected position and with the expected velocity, as discussed in the previous sub-section. In the images from the latter data set (see Fig. 4), we can identify S42, S2 and Sgr A*. However, we do not find any further sources. Given the predicted brightness of S62$_{Peissker}$, it would have to be outside the FOV of either pointing. In 2022,
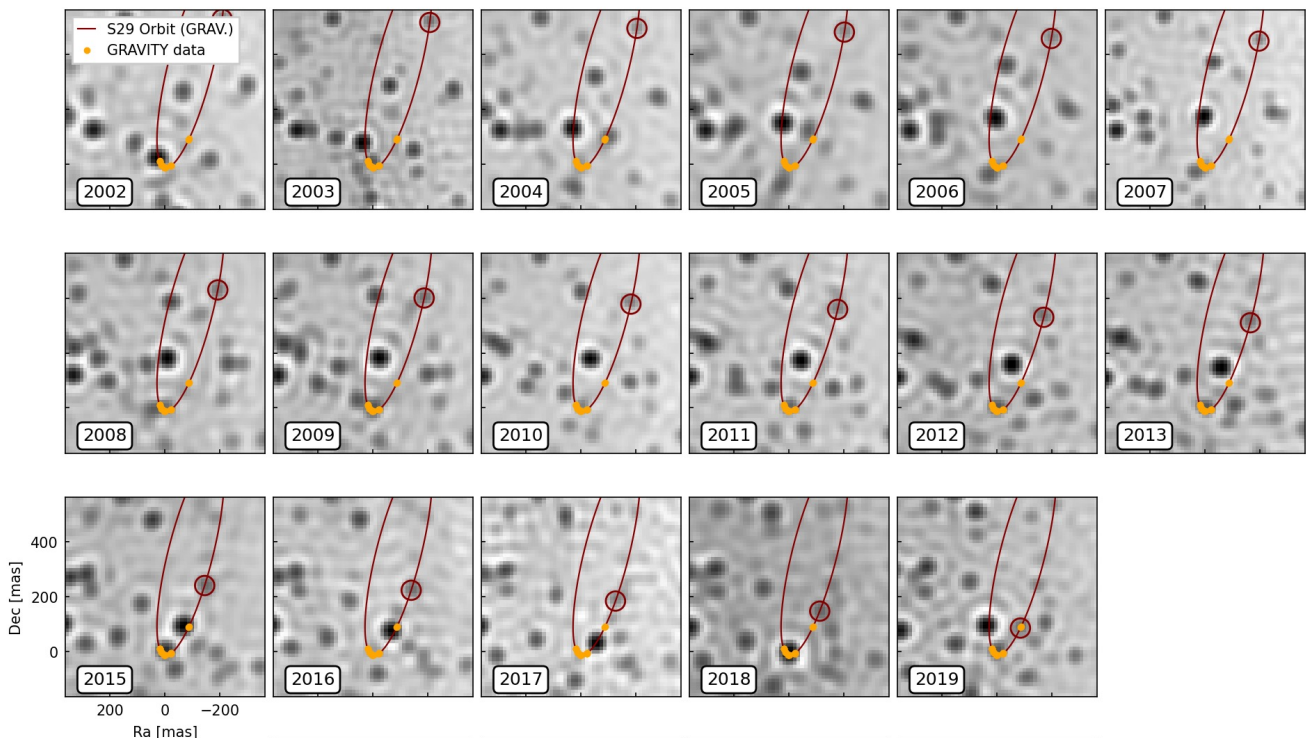
**Fig. D.1.** Compilation of deconvolved NACO images cut to the central ±400,mas. The small circle in dark red marks the position of S29 on its orbit (dark red ellipse) in each panel. The orange data points are the 2019 and 2021 GRAVITY measurements. Evidently, S29 can be traced consistently over two decades. We note that around 2015 - 2017 the star is confused with other stars, such that the astrometry becomes unreliable, but the identification remains unambiguous.

the star is predicted even closer to Sgr A* and should then appear in the central pointing, that is, when positioning the GRAVITY fibers directly onto Sgr A*.

## Appendix E: Probabilistic view of the imaging results

In Figs. 1 and 5, we show the main imaging results of this publication. These images are computed as the mean over all samples in a single $G^R$ run and have been convolved with a Gaussian of 1.6 standard deviation to account for the typical size of the CLEAN beam.

However, the MGVI samples contain information beyond the mean and can also be used to estimate the uncertainty of the result. Since our sky model for the GC is a collection of point sources and the main goal of the analysis is the search for faint, yet unknown stars, we visualize the information contained in the MGVI samples in the following way. In the mean image, we select all pixels whose flux exceeds the background by at least a factor of 10. This threshold is chosen to be below the optimal sensitivity found in Sect. 4.5. Then, for all potential sources identified in this way, we compute the flux variance in the corresponding pixel and express the mean flux in units of its standard deviation. We summarize the position of all potential sources in a RA-Dec plot, as the one shown in Fig. E.1. Thereby, the symbol size indicates the mean flux in units of its standard deviation, that is, the significance of the source candidate, while the color signals its flux. For known sources with a Gaussian position prior, $G^R$ directly infers the coordinates and the standard deviation of this estimate is overdrawn on the source symbol.

This statistical view of the imaging results is provided in Fig. E.1 for the same $G^R$ runs, which are shown in Fig. 1. As

Fig. E.1 demonstrates, S300 corresponds not only to the brightest pixel in the image of faint sources, but it is also the only pixel where flux is detected at a high significance. The only exception here is the July result, where S300 is most strongly affected by fiber damping and $G^R$ fails to detect its flux in a single high-significance pixel. To evaluate the $G^R$ results, we produce and inspect this type of figure for each of the ten $G^R$ imaging runs separately. The results of this inspection are reported in Sect. 4.3.

Further, we also combine the samples from all ten $G^R$ runs into a single figure for a joint representation. These are shown, for all nights considered in Tables 1 and 2, in Figs. E.2 and E.3 respectively. Here, we did not perform a selection of runs, which means that the figures also contain poorly converged runs and runs that failed to identify S300. Furthermore, as we report in Sect. 4.3, the position at which S300 is detected can vary by a single pixel. The same holds for sources in the mosaicing data set, which are not modeled as a point source which is superimposed on the image. Both effects, the inclusion of all runs and the position uncertainty, lead to a decrease in the significance estimate of faint sources in the joint representation. The effect is very apparent from the comparison of Fig. E.1 and Fig. E.2. Also for this reason, the evaluation of the $G^R$ results run-by-run as presented in Sect. 4.3 is a very important step of the analysis.

## Appendix F: Source injection tests

We performed a series of injection tests to further corroborate the result from Sect. 4.3. Injecting a source into the data and trying to recover it in the images has the additional benefit of making it possible to dial down the source's flux step-by-step and in this fashion test our sensitivity.
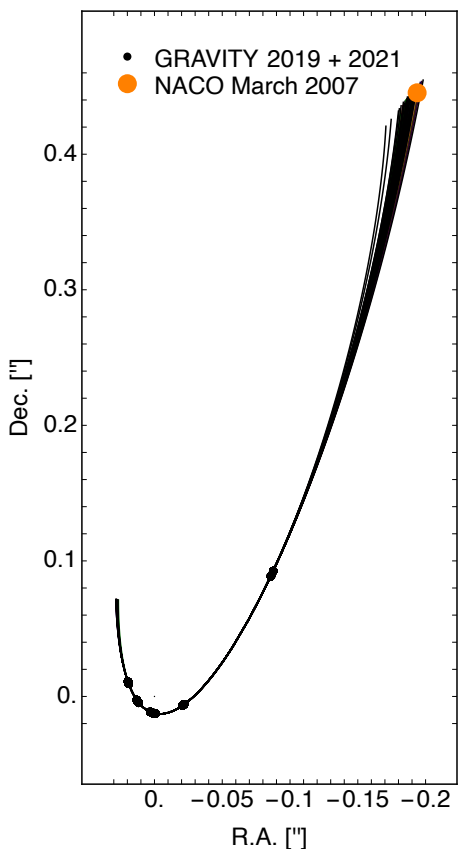
**Fig. D.2.** Backwards prediction of the GRAVITY data for S29$_{GRAVITY}$ leads to a position in March 2007 compatible with the position of S29$_{NACO}$. The black points show the GRAVITY data from 2021 and 2019. The set of black orbits is obtained by bootstrapping these data and fitting each of these mock data sets. The family of orbits naturally matches the position obtained from NACO (orange disk). The epoch March 2007 was chosen since it is the one used in Gillessen et al. (2009) to define the names of the S-stars.

To create the data for the injection tests, we pick the imaging run with the cleanest result from the 2021-05-29 night, this corresponds to the image shown in Fig. 1. The posterior mean gives a model prediction for closure phases and visibility amplitudes, which we subtract from the data to obtain the residuals. We then augment the model by an additional faint source, again compute closures and amplitudes, and, finally, we add back the residuals. In comparison to creating mock data, where we would specify the model by hand and then add a random Gaussian noise realization, this approach preserves the effect of any instrumental systematics as well as correlations between individual data points, which are unaccounted for by the prior model and the likelihood.

In summary, we tested mock sources at two different positions, each with four values for its flux. The first position is at $s_{mock1} = (-15.0, 5.0)$ mas from Sgr A* and thus relatively close to S300. The second location we pick at identical separation from the field center with $s_{mock2} = (5.0, 15.0)$ mas. The flux ratios with respect to S29 for each of these sources are 0.055, 0.037, 0.018, and 0.0037. With $m_K (S29) \simeq 16.6$, this corresponds to a 19.7, 20.2, 21.0 and a 22.7 magnitude mock source, respectively.

The most important results of the injection tests are summarized in Fig F.1. From the first row of plots, that is, the injec-

**Table G.1.** Rms in the central $74 \times 74$ mas of the residual images in Fig. G.1.

| Date | rms/S29 | magnitude |
|------|---------|-----------|
| 2021-03-30 | $1.6 \times 10^{-2}$ | 21.1 |
| 2021-05-29 | $1.2 \times 10^{-2}$ | 21.7 |
| 2021-06-24 | $1.3 \times 10^{-2}$ | 21.6 |
| 2021-07-26 | $2.2 \times 10^{-2}$ | 21.0 |

**Notes.** We convert the flux ratio with respect to S29 into a K-band magnitude, assuming that $m_K (S29) = 16.6$.

tion tests at position 1, close to S300, it is very apparent that the algorithm struggles to properly disentangle both stars. For the highest flux ratio tested, the mock source is found in three runs as a single bright pixel at high significance and in another three smeared out over multiple pixels. Most importantly, however, the scatter around S300 and the amount of flux found in its side lobes increases significantly. The detection of S300 gets cleaner when the flux of the mock source is reduced. But it is already for the second highest flux ratio that the algorithm struggles to robustly detect the mock source and only infers it as flux spread out over multiple pixels at a low significance in six of the ten imaging runs. The result is thus very similar to S300 in the Sgr A*-centered July images. For a flux ratio of 0.018, the mock source at the first position is found only once at a low significance. At this point, it is clear that we have exceeded our sensitivity.

At the position of the second mock source, in contrast, the sensitivity is considerably higher. These results are shown in the second row of Fig. F.1. For a flux ratio of 0.018 or magnitude of 21.0, all imaging runs infer flux at the correct position, five of them in a single pixel with high significance and the remaining spread out over multiple neighboring pixels. The faintest magnitude we tested is 22.7; at this point, only one of the ten imaging runs is able to identify the mock source and we would not be able to claim a detection. Interestingly, however, more flux is attributed to a spurious source in the side lobe west of S55, which is rather close to the injected star.

## Appendix G: Residual images

Residual images play a crucial role for the CLEAN algorithm in that they guide the investigator when deciding where to place the clean box. Furthermore, the highest residual inside this box is understood as a physical source, and its signal is subtracted from the visibilities.

For the G$^R$-code, residual images do not have the same significance. We worked with closure phases and visibility amplitudes and performed the comparison between the model and data directly in the data space by evaluating the likelihood. Indeed, to be able to compute a residual image, it is necessary to invert the Fourier transform in the response equation, namely, Eq. (3), which requires the full phase information of the complex visibilities. Even then, the normalization term, bandwidth smearing and aberration corrections (see Sect. 2.2 and Appendix A) cannot be accounted for.

To maintain an understanding of the spatial structure of our residuals and to aid in the comparison of other methods, here we provide retroactively computed residual images. These make use of the absolute phase measurement which is provided by GRAVITY but was not considered in the imaging.

Closure phases are insensitive to a global translation and our model therefore fixes Sgr A* at the center of the image. In reality,
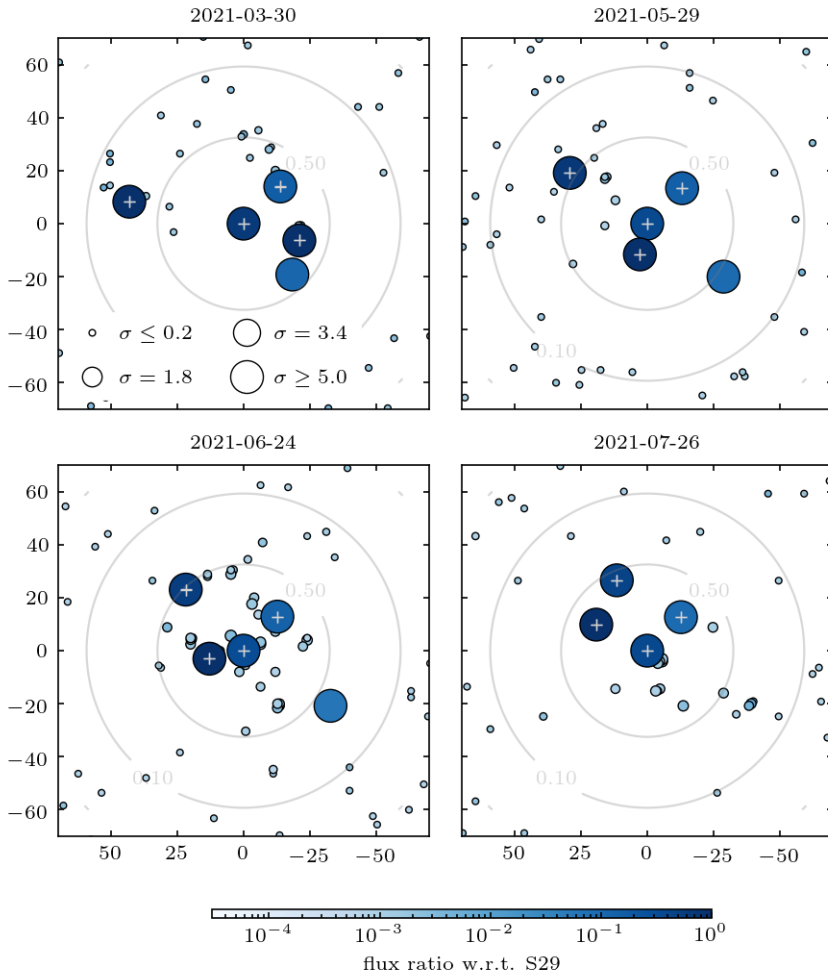
**Fig. E.1.** Statistical view of the imaging results shown in Fig. 1. The symbol color indicates the flux of a source candidate normalized to S29, while its size represents the significance (see Appendix E for details). The Sgr A* flux depicted here has to be multiplied by the light curve at each exposure to arrive at the true flux ratios. For stars which are modeled as a point source, the position uncertainty is indicated in gray.

a slight offset from zero is plausible and would reflect itself in the visibility phases. To correct for this offset, we compute a pseudo-dirty image from our model, and fit a Gaussian to its brightest peak. Doing the same in the dirty image, we can align the model with the data by shifting the coordinates of either Gaussian on top of each other.

We also need to fix the normalization term which is implemented in our model. To determine the overall scaling of the residual images, we use the flux in the brightest pixel of the pseudo-dirty image which corresponds to S29. Our final residual images are given by the inversely Fourier transformed residuals, divided by the S29 signal and are shown in Fig. G.1.

The imaging analysis of GRAVITY GC data with CLEAN in Gravity Collaboration et al. (2021a) used the residual images to estimate the noise level as the root mean square (rms) in the central $74 \times 74$ mas of the image. We provide the same values for all images shown in Fig. G.1 in Table G.1. We note that all images are found to be below the noise level reported in Gravity Collaboration et al. (2021a).

Although a comparison with CLEAN based on the residual images is tempting, we caution the reader against overinterpreting these results. Any telescope-based phase errors are reflected in the residual images, even though $G^R$ is completely insensitive to them. The inversion of the measurement equation to translate the residuals into the image domain cannot account for several important effects such as normalization by the photometric flux, bandwidth smearing, or aberration maps. Finally, the alignment of our model with the absolute position information available

from visibility phases is not perfect and may even have the effect of increasing the noise level by itself.
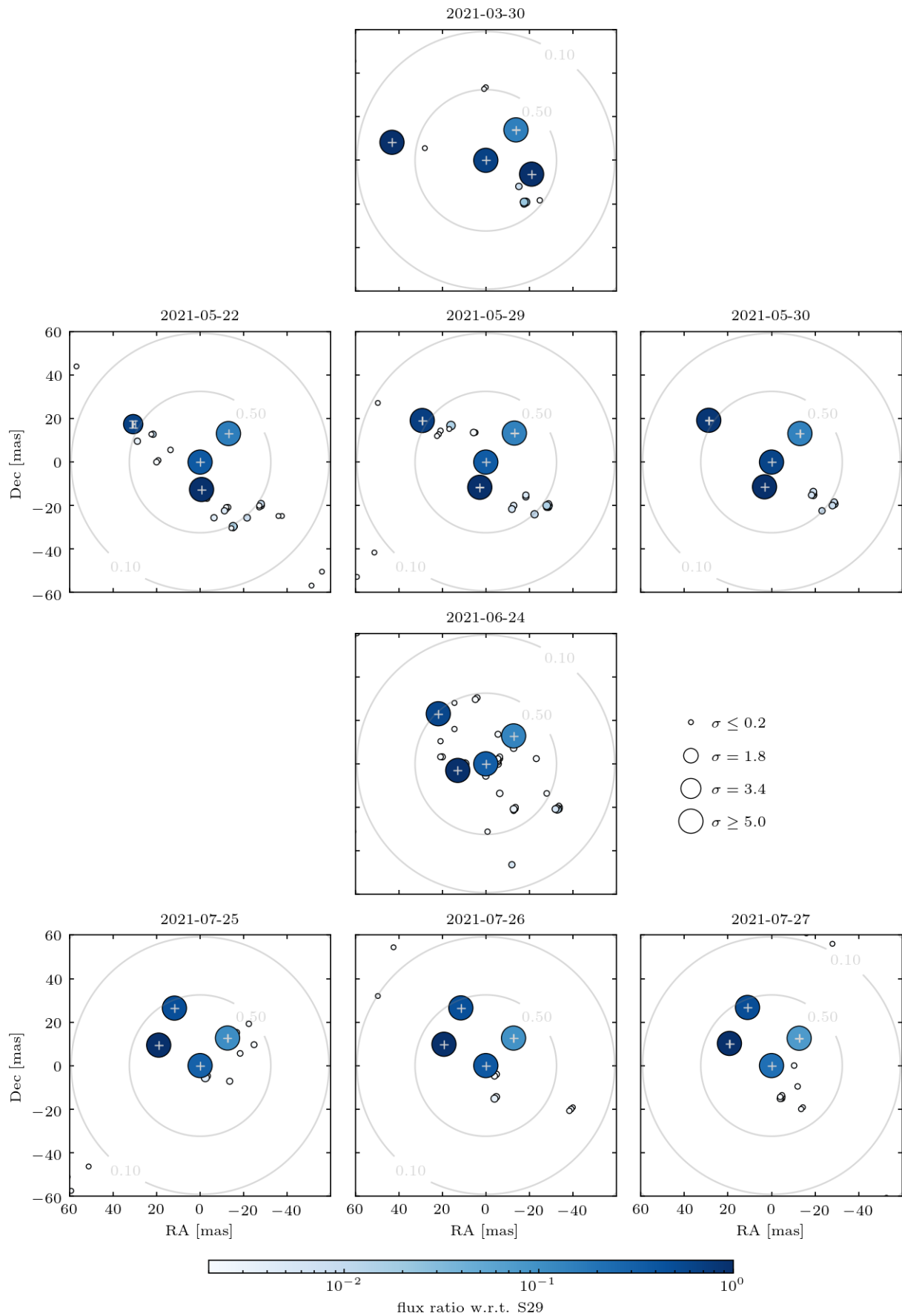
**Fig. E.2.** Summary of all imaging runs for Sgr A*-centered exposures. Each panel combines the samples from ten imaging runs with independent random seeds (see Appendix E for details). The symbol color indicates the flux of a source candidate normalized to S29, while its size represents the significance. The Sgr A* flux depicted here has to be multiplied by the light curve at each exposure to arrive at the true flux ratios. For stars modeled as a point source, the position uncertainty is indicated in gray.
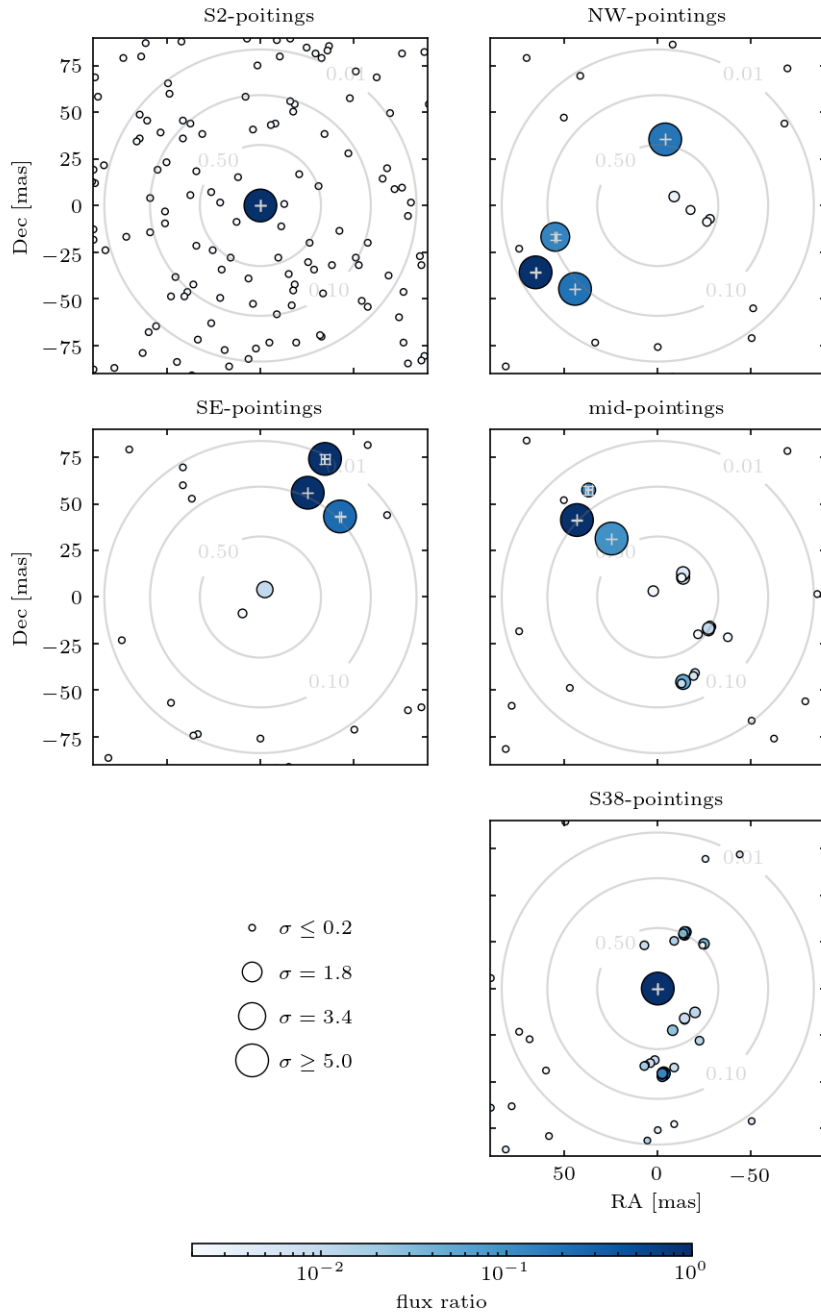
**Fig. E.3.** Summary all imaging runs for the mosaicing data set. Each panel combines the samples from ten imaging runs with independent random seeds (see Appendix E for details). The symbol color indicates the flux of a source candidate, while its size represents the significance. The flux in the images is normalized to S2 in the top left panel, to S38 in the bottom right panel, and to S29 in all other instances. The Sgr A* flux depicted here has to be multiplied by the light curve at each exposure to arrive at the true flux ratios. For stars modeled as a point source, the position uncertainty is indicated in gray.
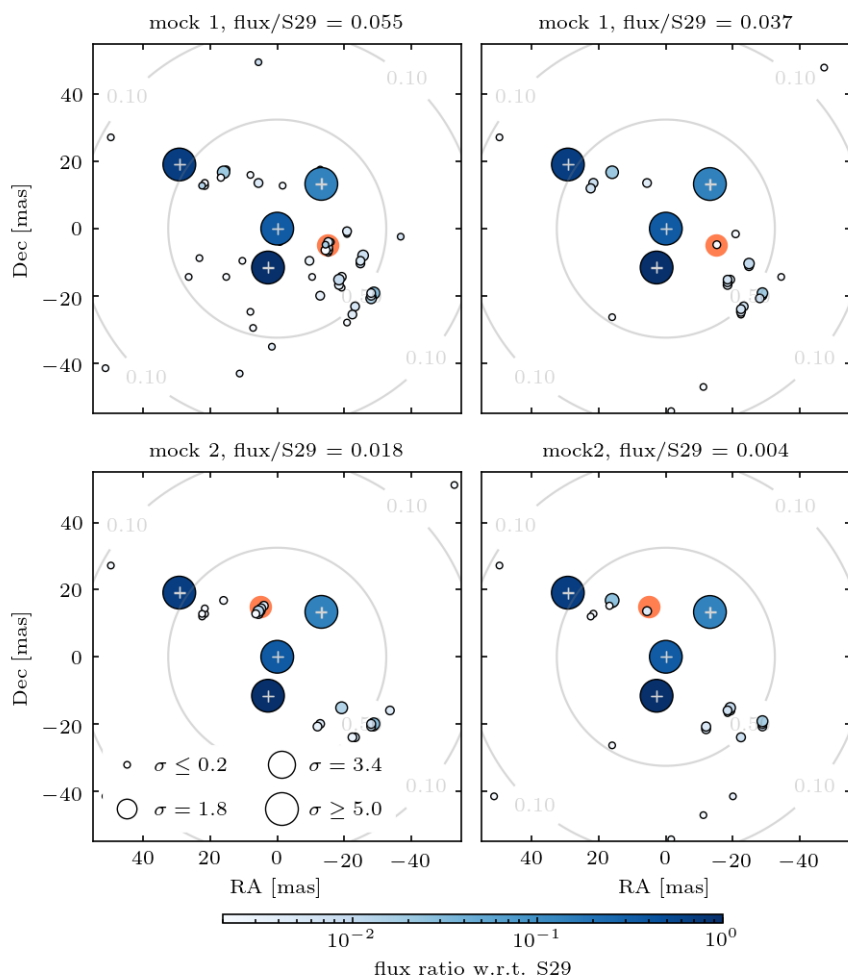
**Fig. F.1.** Results of injection tests for the May 29 data set. The position of the injected source is at $(-15.0, -5.0)$ mas from Sgr A* for mock 1, at $(5.0, 15.0)$ mas for mock 2 and marked by an orange circle in the respective plots. We show the combined samples from ten imaging runs with varying random seed and indicate the mean inferred flux, normalized by S29, by symbol color and the significance of a source candidate by the size of its symbol (see Appendix E for details). In this fashion of display, variation of a faint source's position by a pixel between the imaging runs leads to a decrease in the estimated significance, even if the detection is very robust ($> 5\sigma$) for individual runs. The Sgr A* flux depicted here has to be multiplied by the light curve at each exposure to arrive at the true flux ratios. For stars that are modeled as a point source, the position uncertainty is indicated in gray.
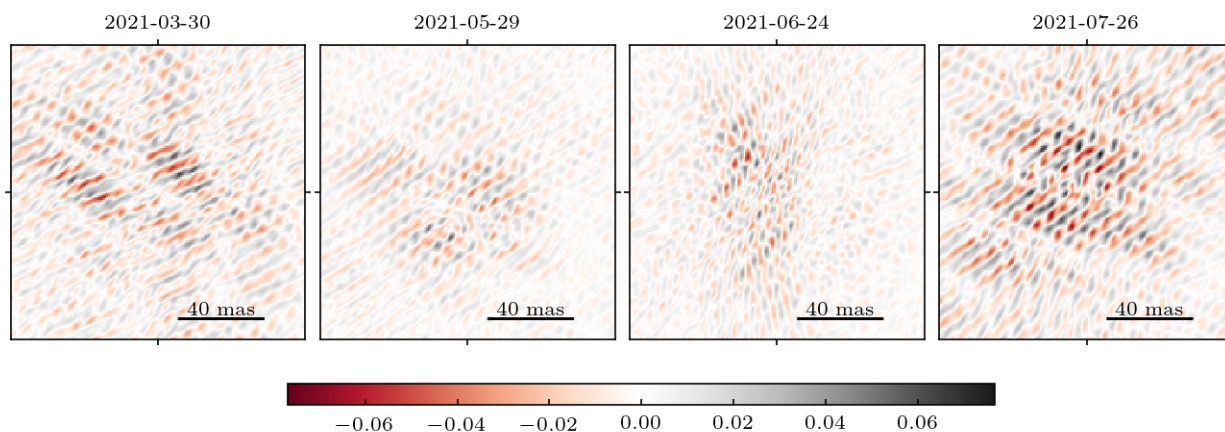


**Fig. G.1.** Residual images for the four observing epochs in 2021, where north is up, east to the left and the flux is normalized to S29. The images corresponding to the nights depicted here are shown in Fig. 1. We note that while our posterior model was optimized on closure phases and visibility amplitudes, here we compare it to the full complex visibility.