The Astrophysical Journal, 925:122 (15pp), 2022 February 1 https://doi.org/10.3847/1538-4357/ac3a6b

Variational Image Feature Extraction for the Event Horizon Telescope

Paul Tiede[1,2,3,4,5], Avery E. Broderick[1,2,3], and Daniel C. M. Palumbo[4,5]
[1] Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON, N2L 2Y5, Canada; paul.tiede@cfa.harvard.edu
[2] Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada
[3] Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, ON, N2L 3G1 Canada
[4] Black Hole Initiative at Harvard University, 20 Garden Street, Cambridge, MA 02138, USA
[5] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
Received 2020 December 6; revised 2021 November 13; accepted 2021 November 15; published 2022 January 31

Abstract

Imaging algorithms form powerful analysis tools for very long baseline interferometry (VLBI) data analysis. However, these tools cannot measure certain image features (e.g., ring diameter) by their nonparametric nature. This is unfortunate since these image features are often related to astrophysically relevant quantities such as black hole mass. This paper details a new general image feature-extraction technique that applies to a wide variety of VLBI image reconstructions called *variational image domain analysis*. Unlike previous tools, variational image domain analysis can be applied to any image reconstruction regardless of its structure. To demonstrate its flexibility, we analyze thousands of reconstructions from previous Event Horizon Telescope synthetic data sets and recover image features such as diameter, orientation, and ellipticity. By measuring these features, our technique can help extract astrophysically relevant quantities such as the mass and orientation of the central black hole in M87.

Unified Astronomy Thesaurus concepts: Radio astronomy (1338); Submillimeter astronomy (1647); Active galactic nuclei (16); Black holes (162); Astronomy software (1855); Computational astronomy (293)

1. Introduction

Generating quantitative measurements about intrinsic radio images from very long baseline interferometry (VLBI) is a computationally and theoretically difficult task. Even in principle, it is impossible to associate a single unique image with a given set of visibility data. In practice, the small number of stations participating in the Event Horizon Telescope (EHT), and thus sparse coverage in the u–v plane, results in a variety of potential image structures. As a result, the process of reaching quantitative conclusions about image features requires significant additional analysis. For the EHT analysis of the 2017 M87 data, this has taken two complementary forms (Event Horizon Telescope Collaboration et al. 2019a).

The first is a traditional Bayesian parametric modeling approach (Event Horizon Telescope Collaboration et al. 2019e, hereafter Paper VI). Simple geometric models are fit to the visibility data, and direct quantitative inferences about the image properties encoded within the model, for example, diameter, are possible. Higher model fidelity improves the accuracy of the derived measurements, and thus this requires an input assumption that the key model features capture the "truth".

The second is a nonparametric approach, usually referred to as "imaging" (Event Horizon Telescope Collaboration et al. 2019d, hereafter Paper IV). This category of methods is broadly defined and includes deconvolution algorithms like CLEAN (Högbom 1974; Schwarz 1978; Clark 1980; Schwab 1984) and forward modeling approaches like "maximum entropy" (Frieden 1972; Gull & Daniell 1978; Narayan & Nityananda 1986), regularized maximum likelihood (RML) analyses (Chael et al. 2016, 2018; Akiyama et al. 2017a, 2017b), and Bayesian imaging (Broderick et al. 2020). The output of imaging by the EHT has been an ensemble of image reconstructions that reproduce the observed visibility data (Paper IV). The imaging methods have the significant advantage that they are extremely flexible and therefore are reasonably expected to cover the "truth". However, unlike parametric modeling, imaging methods do not give direct quantitative measurements of the image features of interest, for example, ring diameter, width, orientation, and so on. Therefore, reaching quantitative conclusions about the image properties requires an additional processing step, which we call "feature extraction". It is this final step that is the subject of this paper.

Feature extraction is similar to the Bayesian parametric modeling applied in the image, rather than visibility, domain. However, it differs in one important respect: by virtue of being performed *after* imaging, the class of applicable "models", that is, image features to be measured, is already well known. For example, in Paper IV and Paper VI, quantitative image features were extracted by the algorithm "ring extractor" `REx` (Chael 2019). However, `REx` is only applicable to images that have a dominant ring-like feature. In general, images from the EHT can have a complex structure and are dependent on the intrinsic source. The active galactic nuclei 3C 279 displays a jet morphology, which, while poorly described by a ring, can be described by a set of Gaussians (Kim et al. 2020).

One possible approach to feature extraction is to "template" relevant image features using a transformation. For example, the Hough transform (Hough 1964; Duda & Hart 1972) is used to extract rings and other shapes from images using template matching. A related method is to approximate the complicated image reconstruction with parametric templates that describe the features of interest. This idea is more akin to the parametric/geometric modeling approach of visibility data in Paper VI and is the approach taken in this paper.

Any comparison requires a suitable quality metric, that is, objective function. Because the total flux is typically an arbitrary rescaling and the image brightness is positive definite,

there is a natural identification between the flux-normalized image and probability distribution. This motivates the use of "divergences" as an extremely flexible class of objective functions for comparing images; for this reason divergences have been used extensively in image processing (e.g., Aherne et al. 1998; Choi & Lee 2003; Goudail et al. 2004).

We adopt the method similar to variational inference (Blei et al. 2017), in which complicated distributions are approximated by simple parametric forms, with parameters estimated via the minimization of an appropriate divergence. In this paper, we develop a number of appropriate parametric forms and explore the performance of a variety of divergences for application to image feature extraction. Therefore, we call this method *variational image domain analysis*, or VIDA. The VIDA algorithm has been implemented in the open source package `VIDA.jl`[6] written in Julia (Bezanson et al. 2017).

There are two main reasons that imaging reconstruction followed by feature extraction is advantageous to directly fitting simple geometric models to the data. First, choosing the correct geometric model is difficult before imaging, meaning imaging is the first step in both methods. Second, fitting simple geometric models can "underfit" the data, leading to biased results. For instance, in Paper VI additional "nuisance" Gaussians were required to obtain a reasonable reduced chi-square. On the other hand, Bayesian imaging techniques (e.g., Broderick et al. 2020) make minimal assumptions about source structure. Combining the Bayesian imaging posterior with VIDA then provides a deterministic map from image to feature posteriors. VIDA's feature posteriors can then be compared to the geometric modeling results that test whether the features are robust across methods. Therefore, VIDA fills a gap in the EHT modeling pipeline and is generic, unlike current EHT tools.

The layout of the paper is as follows: In Section 2, we present the details of VIDA. The different types of templates implemented and the objective function used to find the best approximation to the true image are detailed. Section 3 applies VIDA to a variety of ring-like image reconstructions from the test set of Paper IV and compares the results to REx. This is an empirical demonstration that we can recover the optimal template, even though the objective function is nonconvex. In Section 4, we demonstrate VIDA's flexibility by applying it to nonring images from the test set of Paper IV. Finally, the conclusions are detailed in Section 5. Additional material related to the model and results for CLEAN is presented in Appendices A–C.

2. Variational Image Domain Analysis

The critical insight behind VIDA is that images (sans polarization and modulo total flux) and probability densities are in one-to-one correspondence. Namely, images are point-wise positive and integrable. Probability divergences are a natural class of objective functions used to compare two probability distributions. Furthermore, they have been used in image feature extraction and similarity measures before. Divergences thus form a natural objective function between the image reconstruction and the template. VIDA consists of three ingredients:

1. An **Image** $I(\alpha, \beta)$ whose features we want to extract.
2. A **template** (or approximate image) that parameterizes the features of interest, e.g., ring radius.
3. A **divergence**, i.e., the objective function we minimize.

Each of these building blocks have independent abstract types in `VIDA.jl`, enabling users to easily implement additional images, templates, and divergences.[7] In Sections 2.1 and 2.2 we will review the templates and divergences currently implemented in VIDA respectively.

2.1. Image Templates in `VIDA.jl`

The choice of template used will depend on the structure of the image. For example, the images of M 87 from Paper IV are ring-like, while the reconstructions of 3C 279 from Kim et al. (2020) can be described by several Gaussian brightness distributions. In this section, we present the various templates that are implemented in `VIDA.jl`.

2.1.1. Gaussian Template

To model a source of compact flux we include an asymmetric Gaussian template. The parameters of the Gaussian template are as follows:

1. The size, $\sigma = \sqrt{\sigma_a \sigma_b}$, where $\sigma^2_{a,b}$ are the variances in the principal directions of the Gaussian.
2. The ellipticity, $\tau = 1 - \sigma_b/\sigma_a$, measures the ellipticity of the Gaussian and we assume $\sigma_a > \sigma_b$.
3. ξ, rotation angle (relative to the Gaussian center) of the principal axes measured east of north.
4. x_0, y_0, the center of the Gaussian.

2.1.2. Disk Template

In Section 3 we will test `VIDA.jl` on a number of synthetic data tests. One of the test images is a disk. To approximate disks we use the following template:

$$f_{\rm Disk}(r; r_0, \alpha, x_0, y_0) = N \begin{cases} 1 & r < r_0 \\ \exp(-(r - r_0)^2/2\alpha^2) & r > r_0, \end{cases} \quad (1)$$

where r_0 is the radius of the flat disk, α controls the smoothness of falloff, and N ensures the template is normalized. The radial distance, r, is relative to the center (x_0, y_0). When $r_0 = 0$ this template reduces to a symmetric Gaussian with standard deviation α.

2.1.3. Ring Templates

One of the principal quantities of interest in images of M 87 is the ring diameter, d_0, since it is related to the mass of the central black hole. Additionally, the ring is expected to have some thickness, w, due to the emitting material around the black hole. The simplest template would be a circular Gaussian ring with some thickness. Doppler boosting, however, can cause the emission to appear asymmetric. To model this, a slash can be added to the ring template to include the brightness asymmetry. Additionally, the ring itself does not have to be circular. Ring ellipticity could occur from, for example, the emitting material not being azimuthally symmetric around the black hole. In addition, due to the sparse coverage of the EHT array and imaging algorithms, ellipticity may be introduced

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

[7] See the documentation at https://ptiede.github.io/VIDA.jl/dev/ for a tutorial on how to add additional templates.

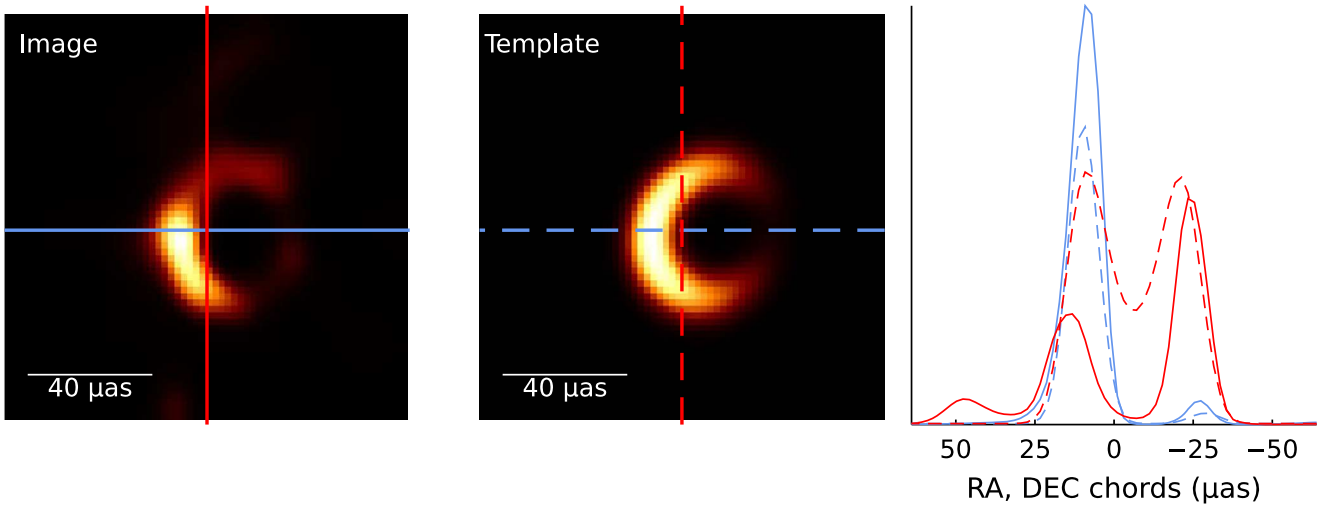

Figure 1. An example of `VIDA` run. Left: image reconstruction of a GRMHD simulation from Paper V and Paper IV. Middle: `VIDA` reconstruction using the `GeneralGaussiantemplate` and Bh divergence. Right: vertical (red) and horizontal (blue) chords through the center of light of the truth image. The dashed lines are for the `VIDA` optimal template, and solid for the reconstruction. This plot can be made in the `VIDA.jl` package using the `triptic` function and is used to assess the quality of the template approximation.

into the reconstructions. Consolidating each of these features into a template parameter, we get the following:

1. d_0: the geometric mean of the semimajor, a, and semiminor, b, axis $d_0 = 2\sqrt{ab}$, which is related to the area of the ellipse, $\pi(d_0/2)^2$.
2. τ: the ellipticity of the ellipse, $\tau = 1 - b/a$.
3. ξ_τ: the position angle of the semimajor axis measured north of east.
4. w: the width of the Gaussian ring, defined to be the FWHM of the Gaussian, i.e., $w = 2\sqrt{2\log 2}\,\sigma$, where σ is the standard deviation.
5. s: the strength of the slash described in Equation (4).
6. ξ_s: the position angle of the slash measured east of north.
7. (x_0, y_0): the center of the ring.

The functional form of the template is given by

$$h_\theta(x, y) = S(x, y; s, \xi)\exp\left[-\frac{(d_\theta(x, y))^2}{2\sigma^2}\right], \tag{2}$$

where $S(x, y; s, \xi)$ is the slash function and $d_\theta(x, y)$ is the minimum distance between the ellipse with parameters $\theta = d_0$, τ, x_0, y_0, and the point x, y. If $\tau = 0$, $d_{\tau=0}(x, y) = |\,||\mathbf{x} - \mathbf{x}_0|| - d_0/2|$. However, for an ellipse there is no analytical equation and instead one has to numerically minimize the function,

$$L(x, y, e_x, e_y) = ||(x - e_x, y - e_y)||, \tag{3}$$

subject to the constraint that e_x, e_y are points on the ellipse with parameters d_0, τ, ξ_τ.

For the slash function S, we use a first-order cosine expansion in azimuthal angle ϕ around the center x_0, y_0:

$$S(x, y; s, \xi_s) = N_0[1 + s\cos(\phi - \xi_s)], \tag{4}$$

where N_0 is a normalization factor to ensure the template is unit normalized. To prevent image brightness from becoming negative, we restrict $s \in [0, 1]$. In the `VIDA.jl` package, this template is called `GeneralGaussianRing` (GGR), and an example reconstruction using said template is shown in Figure 1.

Additionally, `VIDA.jl` has a number of other ring-like templates currently implemented:

1. `GaussianRing`: symmetric Gaussian ring with constant azimuthal intensity (i.e., GGR with τ, $s = 0$).
2. `SlashedGaussianRing`: symmetric Gaussian ring with azimuthal slash described by Equation (4) (i.e., GGR with $\tau = 0$).
3. `EllipticalGaussianRing`: elliptical Gaussian ring with constant azimuthal brightness (i.e., GGR with $s = 0$).
4. `TIDAGaussianRing`: GGR template where the slash and ellipticity position angle are either aligned or antialigned.

We also include a more general version of the GGR called the `CosineRing{N,M}`. This template is similar to the GGR template but where the width, σ, and slash function (4) are replaced by a higher-order cosine expansion in azimuthal angle ϕ:

$$S_M(\phi; \mathbf{s}, \boldsymbol{\xi}^{(s)}) = 1 - \sum_{m=1}^{M} s_m \cos[m(\phi - \xi_m^{(s)})], \tag{5}$$

$$\sigma_N(\phi; \boldsymbol{\sigma}, \boldsymbol{\xi}^{(\sigma)}) = \sigma_0 + \sum_{n=1}^{N} \sigma_n \cos[n(\phi - \xi_n^{(\sigma)})], \tag{6}$$

where $\boldsymbol{s}$, $\boldsymbol{\sigma}$, $\boldsymbol{\xi}^{(\cdot)}$, are vectors with the cosine expansion coefficients of the slash, standard deviation, and angular offset.[8] We can reproduce the GGR template by setting $M = 1$ and $N = 0$ in Equations (5) and (6), respectively. This template can be used if the image has a ring-like feature that has a bumpy azimuthal profile.

2.1.4. Constant Template

We found that many image reconstructions had a diffuse intensity throughout the image due to poor dynamic range from sparse coverage and regularization effects. To model the background, we added a constant intensity template. This

[8] In this paper, bold characters will represent vectors.

template is typically required to be included in any analysis to ensure reliable feature extraction.

2.1.5. Composite Templates

A general image reconstruction from a VLBI observation may have multiple image features. As such, `VIDA.jl` allows the user to combine multiple features into composite templates, where each individual component is given a relative flux.[9] An example of this is shown in Section 4.1, where three Gaussian templates are used to model the reconstructed image.

2.2. Probability Divergences

As mentioned above, `VIDA.jl` uses an analogy between images and 2D probability distributions to motivate the use of divergences as objective functions. Divergences form a general way to measure the similarity between two distributions. A divergence can be thought of as a functional $F_q[p] = \mathcal{D}(p||q)$, comparing the distribution p (template) to a reference q (image), and is required to be nonnegative, $\mathcal{D}(p||q) \geqslant 0$, and nondegenerate, $\mathcal{D}(p||q) = 0$, if and only if $p = q$. Note that this definition is more general than a metric. Namely, a divergence does not have to be symmetric, that is, $\mathcal{D}(p||q) \neq \mathcal{D}(q||p)$, or satisfy the triangle equality.

One of the most well-known divergences is the Kullback–Leiber (KL) divergence (Kullback & Leibler 1951) or relative entropy:

$$KL(p||q) = \int p(\boldsymbol{x}) \log\left(\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})}\right) \mathrm{d}^2\boldsymbol{x}. \tag{7}$$

One issue with the KL divergence is its definition when the support of q and p differ, that is, if $q(x) = 0$. In this case we set the contribution to the integral to be zero.

In addition to the KL divergence, `VIDA.jl` also includes the Bhattacharyya divergence (Bh; Bhattacharyya 1943):

$$Bh(p||q) = -\log \int \sqrt{p(\boldsymbol{x}) q(\boldsymbol{x})}\, \mathrm{d}^2\boldsymbol{x}. \tag{8}$$

The Bh divergence is related to a well-known metric on probability spaces, the square Hellinger distance:

$$H(p, q) = \frac{1}{2} \int (\sqrt{p(\boldsymbol{x})} - \sqrt{q(\boldsymbol{x})})^2 \mathrm{d}^2\boldsymbol{x} = 1 - Bh(p||q). \tag{9}$$

Therefore, minimizing the Bh divergence is simply a least-squares fit in the space of the square root of distributions.

In this paper we will present the results from optimizing the Bh divergence for two reasons. First, we found that while the KL and Bh divergence produce near-identical results, the Bh divergence required ∼25% less evaluations to converge. Second, the Bh divergence has preferable theoretical properties compared with the KL divergence. Namely, it is well defined when the image pixels have zero intensity and is symmetric.

2.3. Optimizing the Divergence

A problem when using probability divergences is that they give nonconvex, nonlinear optimization problems. Furthermore, the nature of the problem will change if the template changes, making an analytic analysis difficult. Therefore, to extract the globally optimal template, we turned to heuristic global optimizers, such as genetic/evolutionary strategies (see Das & Suganthan 2011, for a review) and simulated annealing (Goffe 1996). For this paper we used the Julia package BlackBoxOptim.jl.[10,11] BlackBoxOptim.jl uses natural evolution and differential evolution strategies to perform a stochastic search of the parameter space. In the next section we will validate that our chosen optimizer is able to recover the optimal template reliably.

3. Validating `VIDA`

To validate `VIDA` we need to analyze two related quantities. First, we need to verify that the objective function, that is, the Bh divergence Equation (8), is robust to the artifacts that occur in image reconstructions, namely, whether the recovered parameter distribution contains the true value. Additionally, given the complex nature of the optimization problem, we need to ensure that the chosen optimizer can recover the global minimum of the divergence. Our validation procedure will consist of the following:

1. Selecting an applicable ground-truth image $I_{\rm truth}$ (see the left column of Figure 2).
2. For each truth image, creating a simulated EHT observation matching the observation characteristic of the EHT M 87 2017 observations.
3. Creating an ensemble of image reconstructions of the simulated observations using the same procedure as Paper IV.
4. Applying VIDA to each image reconstruction and comparing the inferred results to the ground-truth parameters and the reconstruction technique `REx` used in Paper IV and Paper VI.

3.1. Step 1: Selecting Ground-truth Images

To validate `VIDA`, we applied it to a subset of the test set from Paper IV. The sources we considered are shown in Figure 2 and consist of two geometric crescents and a general-relativistic-magnetohydrodynamical (GRMHD) simulation from Event Horizon Telescope Collaboration et al. (2019f). The geometric crescent model is described by

$$I(r, \theta) = F_0(1 - s\cos(\theta - \xi)) \frac{\delta(r - r_0)}{2\pi r_0}. \tag{10}$$

The infinitely thin ring is then convolved with a circular Gaussian with FWHM 10 μas. Two orientations (measured east of north) $\xi = 180°$ and $\xi = 150°$ are considered in this paper and are shown in the top and left panels of Figure 2. For both orientations we took $r_0 = 22$ μas, $s = 0.46$, and $F_0 = 0.6$ Jy. After blurring the ring, it is important to note that the effective radius (the intensity peak) is smaller than the original radius of the ring (see Paper IV). The amount the diameter is biased inward is given approximately by

$$d_{\rm blur} = d_{\rm true} - \frac{1}{8\ln 2} \frac{\alpha^2}{d_{\rm true}}, \tag{11}$$

[9] Note the absolute or total flux of the image is not recoverable since we renormalize each image to have unit flux.

[10] https://github.com/robertfeldt/BlackBoxOptim.jl

[11] `VIDA.jl` also has interfaces to other Julia optimization packages, such as Optim.jl and CMAESEvolutionStategy.jl.

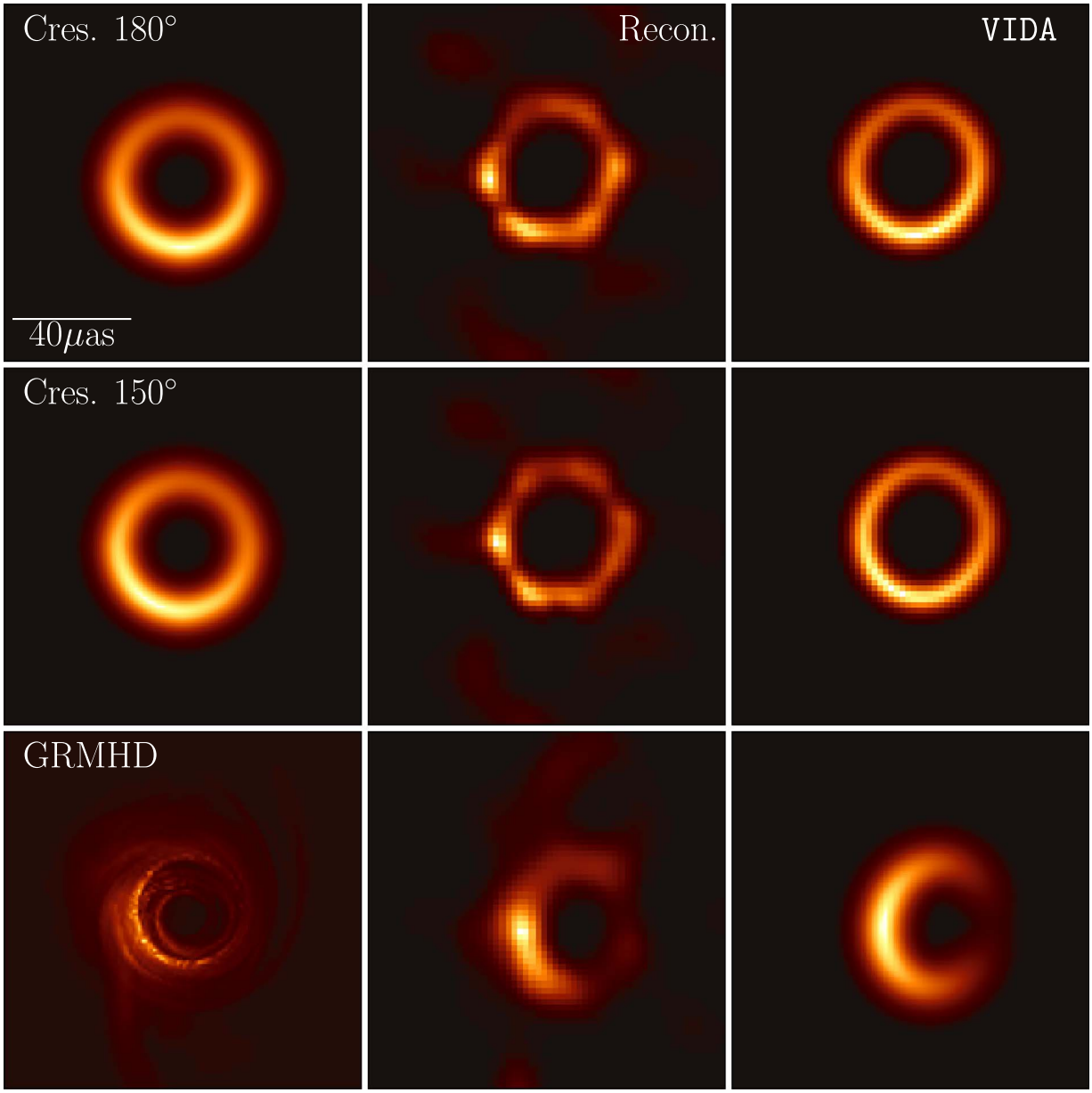

Figure 2. Images used for the imaging validation from Paper IV. We considered three models. Top row: crescent with position angle $\xi = 180°$ north of east; middle row: crescent with position angle $\xi = 150°$; bottom row: GRMHD simulation. The left column shows the truth image, the middle column shows an example reconstruction, and right column shows the optimal VIDA template applied to the reconstruction.

where $d_{\rm true}$ is the diameter of the nonconvolved ring (44 μas) and α is the FWHM of the Gaussian kernel ($\alpha = 10$ μas). Using Equation (11) with $d_{\rm true} = 44$ μas gives $d_{\rm blur} \approx 43$ μas. If we also consider the finite resolution of the EHT array (∼20 μas), this is further decreased to ≈42 μas. Therefore, we expect both VIDA and REx to recover a diameter of 42 μas. Additionally, the slash strength is also modified by the convolution. Fitting the crescent with the GGR template we find $s = 0.32$, which is the value we will take as the ground truth below.

3.2. Step 2: Creating Simulated EHT Observations

While VIDA could be applied to the ground-truth images shown in the left column of Figure 2, this is not applicable to what the EHT observes. The EHT is a very long baseline inferometer and instead observes complex visibilties, $V(u, v)$, which are related to the on-sky image through the van Cittert-Zernike theorem (Thompson et al. 2017):

$$V(u, v) = \int e^{-2\pi i(u\alpha + v\beta)} I(\alpha, \beta) \mathrm{d}\alpha \mathrm{d}\beta. \tag{12}$$

In addition, atmospheric and telescope effects can further corrupt the signal. To model these corruption effects we use the eht-imaging package (Chael et al. 2016, 2018) to generate realistic simulated data.

3.3. Step 3: Generating an Ensemble of Reconstructed Images from Simulated VLBI Data

To validate VIDA we used image reconstructions from the forward modeling or "regularized maximum likelihood methods"

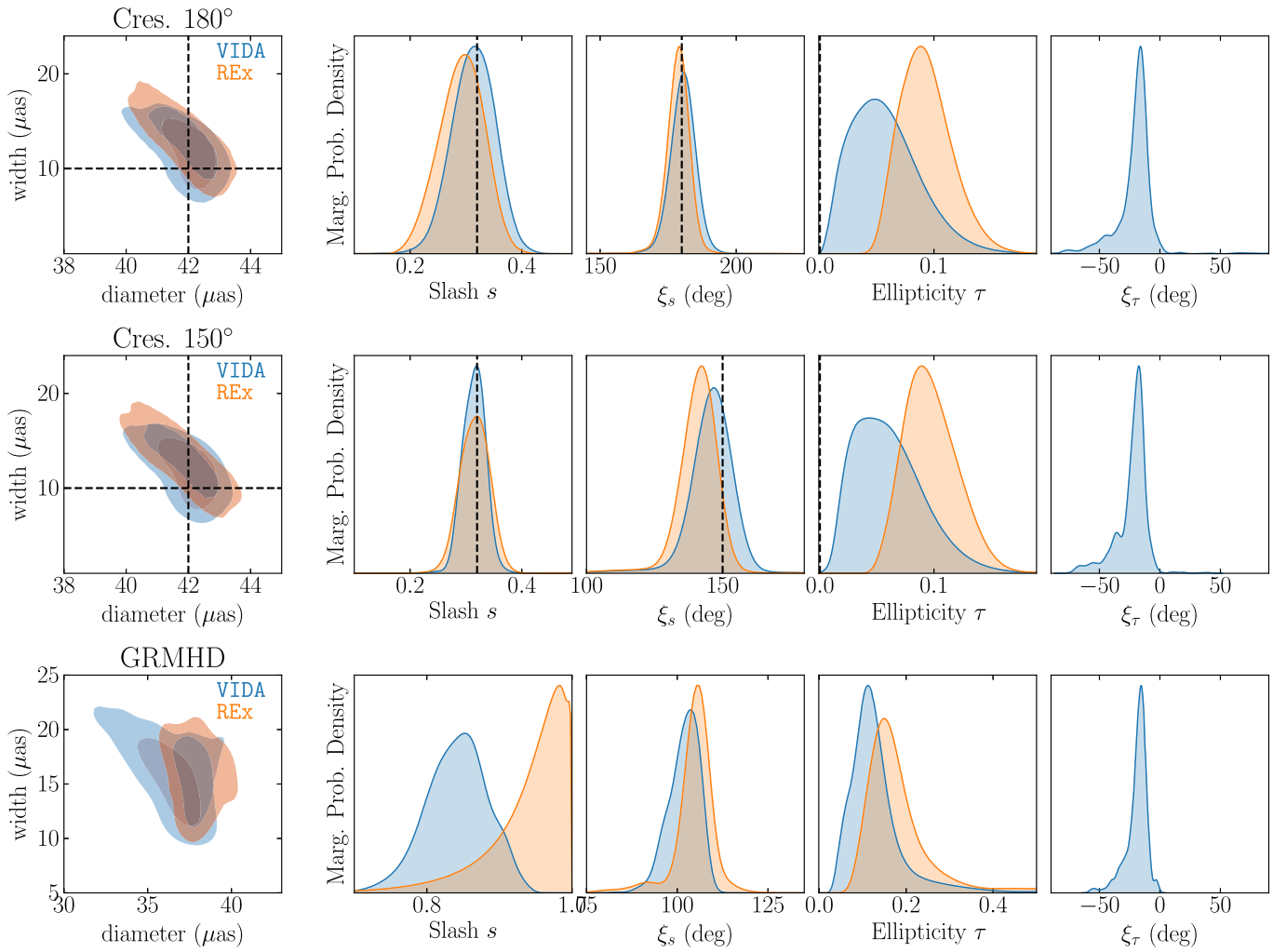

Figure 3. Results from VIDA (blue) and REx (orange) being applied to the Paper IV reconstructions using the simulated April 11 EHT sampling of Figure 2. The top and middle rows show the geometric crescent results, and the bottom row shows the GRMHD. For the crescent models we found good agreement between VIDA (blue) and REx (orange) for the ring diameter, width, slash, and location of the azimuthal brightness. Furthermore, the crescent distributions for the diameter (after accounting for Equation (11)), width, slash strength s, and position angle ξ_s are consistent with the truth image (black dashed line). For the ellipticity, τ, we found VIDA's value was systematically lower than REx's as expected since REx measures the fractional dispersion of the ring radius and random fluctuations can then create an ellipticity floor around 0.05 (see Appendix A for a detailed discussion). The ellipticity orientation, ξ_τ, is only recovered by VIDA, so there is no REx comparison. For the GRMHD simulations, the agreement between REx and VIDA is weaker. Namely, the distributions for the parameters are sometimes significantly different, although the portions of the distributions do overlap. Similar results were found using the simulated EHT coverage from April 5, 6, and 10.

(RML), for example, Honma et al. (2014), Bouman et al. (2016), Akiyama et al. (2017a), Akiyama et al. (2017b), Ikeda et al. (2016), and Kuramochi et al. (2018) and more specifically the `eht-imaging` package (Chael et al. 2016, 2018). The goal of RML methods is to find the image, I, that minimizes the objective function:

$$J(I) = \sum_{\text{data}} \alpha_d \chi^2_d(I) - \sum_{\text{regularizers}} \beta_r S_r(I). \quad (13)$$

Following Paper IV, each χ^2_d is defined solely from the data products from the EHT telescope, for example, complex visibilities. The second term encapsulates our additional assumptions, or regularizers, that are placed on the image. The $\alpha_d\,\beta_r$, are the "hyperparameters" that control the relative weighting of the regularizers and data products. For the list of the regularizers used, see Paper IV. In an attempt to model the uncertainty in the image reconstructions, we used the same set of imaging hyperparameters as Paper IV. The resulting set of image reconstructions is called the "top set" and results in 1572 reconstructions per data set.

For completeness we also apply VIDA to the CLEAN M 87 reconstructions for the Crescent 180, GRMHD, double, and disk in Appendix C. Note that for M 87 there are only 30 reconstructions for each simulated data set, making the comparison more uncertain.

3.4. Step 4: Applying VIDA to the Ring-like Image Reconstruction Ensembles

VIDA was run on each set of image reconstruction ensembles using the GGR template with a constant brightness background whose relative flux was a free parameter, giving nine parameters in total. Some example reconstructions and corresponding optimal templates are shown in Figure 2. The results are shown in Figure 3. For the crescent models, we were able to recover the expected diameter d, width w, and azimuthal orientation ξ_s (black dotted lines in Figure 3). In addition, we

compared the VIDA results to the REx method used in Paper VI.

REx assumes that a single ring-like feature dominates the image reconstruction and then finds the ring by finding the image location that leads to a ring with minimal radial dispersion. REx characterizes (see Appendix A and Paper IV for definitions) the ring through diameter d, width w, brightness moments s, and orientation ξ_s and a fractional dispersion of the diameter f_d. The diameter and width and brightness profile of the REx measurement are similar to VIDA's measurement with the GGR template. However, the fractional dispersion is not directly measured. Instead, VIDA measures the ellipticity of the ring. In Appendix A we demonstrate how f_d and τ are related if the dominant source of radial dispersion in the ring is due to ellipticity.

The agreement between REx and VIDA is excellent for both the crescent images. The peak and overall width of the distribution for each parameter in Figure 3 are consistent between REx and VIDA. The ground-truth values (black vertical lines) for the diameter, width, and brightness orientation ξ_s are also consistent with the REx and VIDA results. We also analyzed the pairwise linear correlations between all the parameters and found that only the diameter and width were correlated (see Figure 9). This correlation is expected from Equation (11) and occurs due to the finite resolution of the EHT array.

To compare VIDA's and REx's measurements of ellipticity, we first note that an additional processing step is needed since the two definitions differ. REx does not directly measure τ but instead measures the fractional diameter dispersion of the ring f_d (see Equation (A5) for a definition). If we assume that the ring's ellipticity dominates f_d, then f_d and τ are related by an invertible map. For more information about this conversion, see Appendix A. In Figure 3 we show REx's results after converting from f_d to τ. Comparing the two measurements of τ, we see that REx's measurement is consistently greater than VIDA's. This bias is not unexpected given that when τ is small, the conversion described in Appendix A no longer applies. Instead, the fractional dispersion is dominated by random fluctuations in the ring diameter, creating a floor in f_d. If we then naively apply the previous conversion, as was done in Figure 3, we will overestimate τ (see Figure 8).

VIDA also recovers the orientation of the ring ellipticity ξ_τ. Interestingly, in all instances, we measure a similar distribution for ξ_τ irrespective of the intrinsic image. This distribution is biased so that the semimajor axis of the ellipticity is in the north–south direction. By visually inspecting the image reconstruction ensembles (see Figure 2 for a typical example reconstruction), we confirmed this feature was present in the reconstructions and was not a bias from VIDA. We also find a nonlinear correlation between the measured ring ellipticity and orientation. Namely, τ is significantly larger when $\xi_\tau \approx 0$. The origin of this bias and implications for M 87 will be discussed in Tiede et al. (2020).

4. Applying VIDA to Additional Simulated Image Reconstructions

In the previous section, we saw that VIDA and REx gave remarkably similar answers to problems that demonstrated similar ring-like structures. While REx is limited to ring extraction, VIDA can be applied to any image given a suitable template function. This section will explore VIDA's capabilities of extracting features from a broader range of potential sources. To accomplish this, we will consider the other nonring test images from Paper IV: the symmetric disk and double Gaussian (see Figure 4). We will follow the same steps in the previous section to evaluate VIDA's performance.

4.1. Double Gaussian

Here we consider a source composed of a compact double with two circular Gaussian components. Each Gaussian has an FWHM of 20 μas. One of the Gaussians is placed at the origin and has a flux of 0.27 Jy, which we will call the NW component. The other Gaussian is at $\Delta\mathrm{RA} = 30$ μas and $\Delta\mathrm{DEC} = -12$ μas and has a flux of 0.33 Jy and will be called the SE component. This type of source could arise when looking at active galactic nuclei using VLBI images, such as the recent 3C 279 results (Kim et al. 2020).

To extract the reconstruction's compact components, we used a template with three asymmetric Gaussian components and constant background. Two of the Gaussian components were allowed to be arbitrary, while the third Gaussian component was forced to be large ($r_0 > 15$ μas). The reason for the third Gaussian was that there tended to be a region of additional emission around the two dominant Gaussian components in the image reconstructions. This diffuse emission can be seen in the top middle panel of Figure 4. If we did not include this third component, we found that the Gaussian components tended to be quite large to soak up the extra emission.

VIDA's results for the double Gaussian are shown in Figure 5. Overall the on-sky size of each Gaussian, their separation, and flux ratio are contained in the support of the distribution. However, there do exist biases in the results, such as the FWHM of the Gaussian being biased low.[12] Additionally, the ellipticity τ, is larger than zero; however, this is to be expected as asymmetry can only be added to the images.

For the simulated EHT sampling from April 6, we see that the ellipticity appears to be bimodal, and the parameter uncertainties are greater than the other days. This uncertainty was unexpected given the relatively good EHT coverage on April 6. After analyzing the image reconstructions we deemed this not to be due to VIDA. Instead, a subset (10%–15%) of image reconstructions on April 6 exhibits a third bright Gaussian feature. If we remove these reconstructions, we find that the results for the simulated April 6 coverage are consistent with the other days.

4.2. Disk Image

The intrinsic image is a symmetric flat disk with a diameter of 70 μas, which is then convolved with a Gaussian with an FWHM of 10 μas. The true image, an example reconstruction, and optimal template for that reconstruction are shown in the lower left, middle, and right panels, respectively, of Figure 4.

To encode the diameter of the disk we use the FWHM of the disk template:

$$d_{\mathrm{disk}} = 2r_0 + 2\sqrt{2\log 2}\,\alpha, \tag{14}$$

[12] Note that the FWHM of the Gaussian is related to the hyperparameters of the top set.

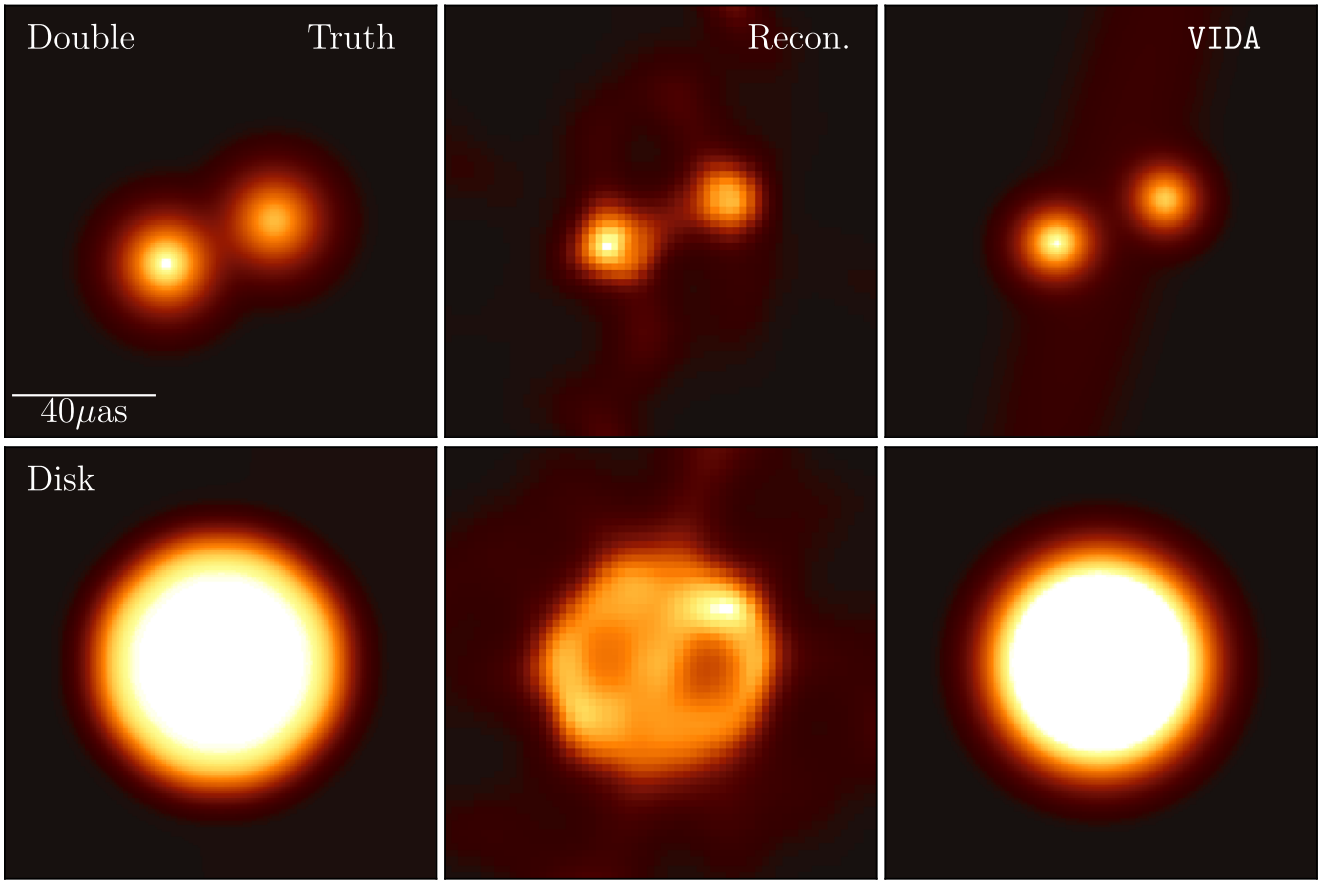

Figure 4. Results from applying VIDA to the nonring test images from Paper IV where the left column shows the ground-truth image, the middle column shows an example reconstruction from the top set, and the right column shows the optimal template from applying VIDA to the reconstruction. Top: results for the double image. Bottom: results for the disk. Overall, VIDA can recover the intrinsic structure of both images.

where r_0 and α are described in Equation (1). We fit the disk template to the ground-truth image to calibrate the diameter definition to the disk's true diameter. We found that the optimal template for the true image had a $d_{\rm disk} \approx 69\,\mu$as. If we convolved the image by an additional $20\,\mu$as to take into account the finite resolution of the EHT array, we found $d_{\rm disk} \approx 68\,\mu$as. This is the value we use as the ground-truth diameter in all comparisons below.

Figure 6 displays the results for VIDA applied to each day. Ignoring the simulated EHT sampling of April 10, which has poor coverage compared with the other days, we find that the results are very consistent between days. For the simulated April 5 EHT sampling, we find the median diameter $d_{\rm disk} = 65.7^{+0.42}_{-0.76}\,\mu$as where the range is the 68% interval about the median. Similarly, for the simulated EHT sampling of April 6, we find $d_{\rm disk} = 65.6^{+0.40}_{-0.47}\,\mu$as, and for April 11, $d_{\rm disk} = 65.5^{+0.45}_{-0.80}\,\mu$as. This demonstrates that VIDA is robust to the slight difference in image reconstructions from different baseline coverage.

For the simulated EHT sampling of April 10, we had a different result finding a bimodal diameter. Analyzing the reason for this, we found that images with $d_{\rm disk} \approx 60\,\mu$as had a markedly different structure than the rest of the images. Given the distinct nondisk structure of the image, it is no surprise that VIDA struggles at recovering the correct diameter.

Comparing our result for the diameter to the true value, $68\,\mu$as, we find that our result has a consistent bias of $\approx 2.4\,\mu$as for the simulated April 5, 6, and 11 EHT sampling. Again this appears to be an artifact of the imaging process. In Figure 7, the radial profiles of the truth (dotted lines), averaged image reconstructions,[13] and optimal templates are compared. VIDA does an excellent job of recovering the size of the images, which are similarly biased toward smaller radii. This suggests that the diameter bias is intrinsic to the top set used for this disk. Furthermore, as shown in Appendix C, the CLEAN reconstructions of the disk do not suffer from the same bias. Given this, it is unlikely that observed bias is due to the finite resolution of the EHT array and is intrinsic to the top set used.

5. Summary and Conclusions

We present VIDA, a new image feature-extraction technique appropriate for use by the EHT. VIDA adopts a forward modeling approach to extract quantitative image properties by approximating the image with a parameterized family of functions that encode the desired image properties.

A key feature of VIDA is its flexibility. Multiple image components have already been implemented, from which composite models of significant complexity can be constructed. These include ring-like templates of particular relevance to EHT images, Gaussians, and constants.

[13] When averaging we first centered the images by computing the image centroid and normalized the images to have unit flux.

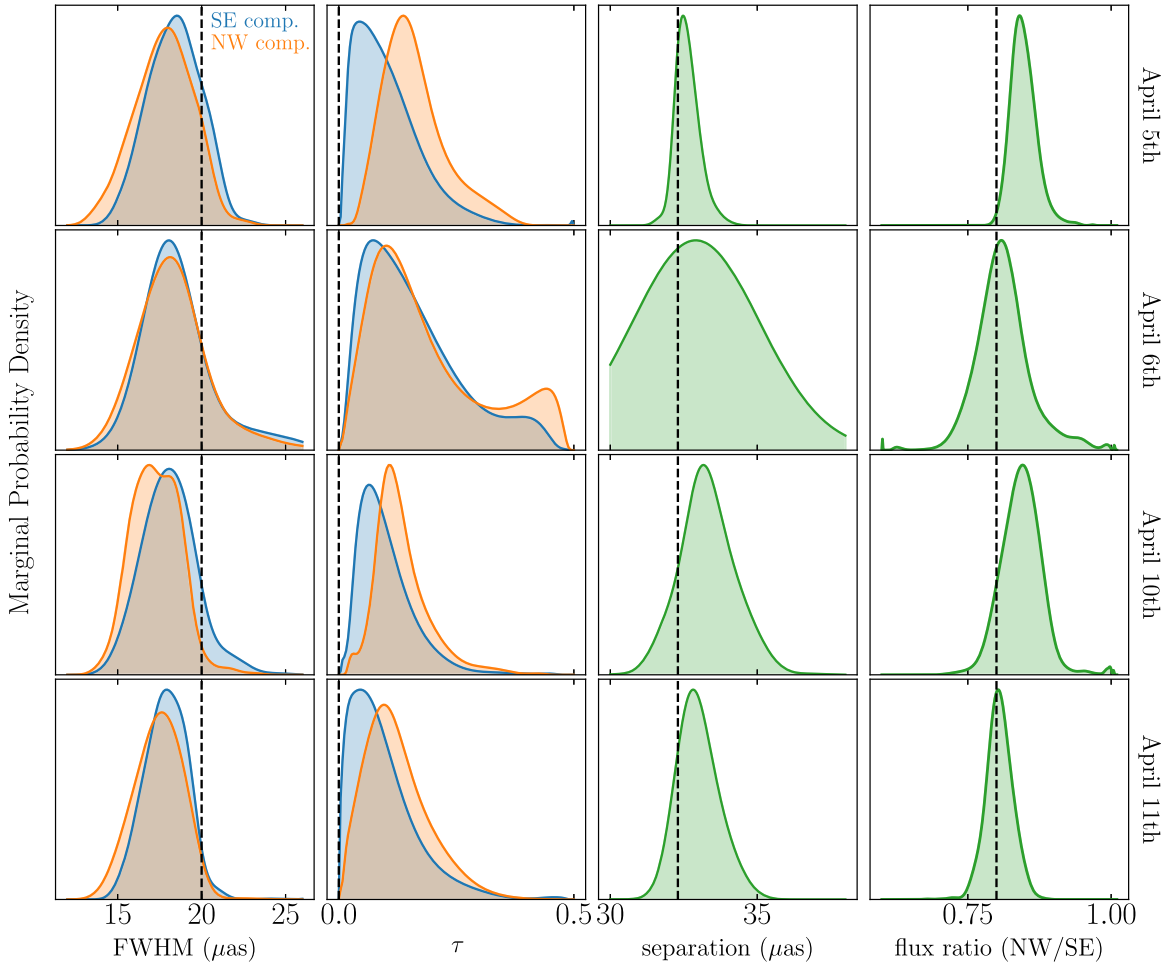

Figure 5. VIDA results for the two compact Gaussian components (blue for SE component and orange for the NW) in the double Gaussian test image. Each row corresponds to a different EHT sampling from the 2017 M 87 observations. The green curves are for parameters that are a combination of the SE and NW components. On all days the true values are included in the support of the parameter distributions found by VIDA. Note that the broad distribution found using the simulated EHT April 6 sampling is due to an imaging artifact as discussed in the paper.

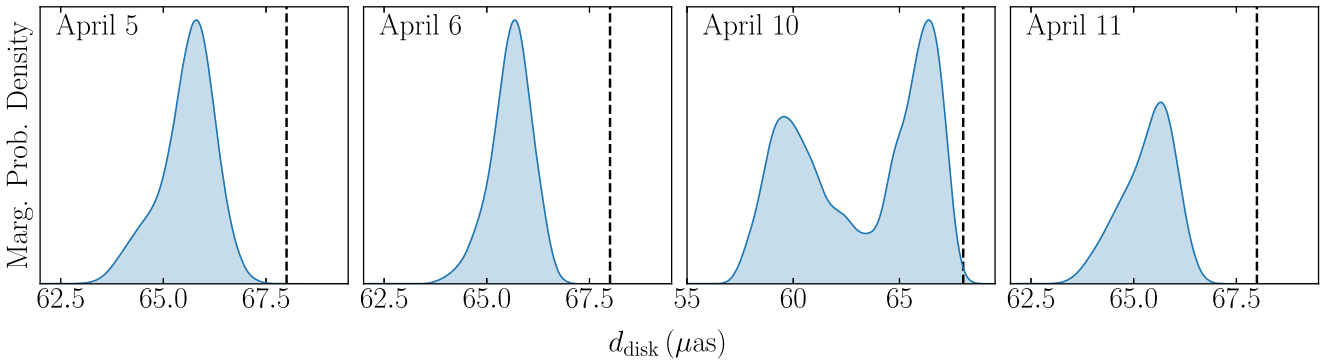

Figure 6. VIDA results of the diameter for the disk top set. The diameter is given by Equation (14). Ignoring the simulated EHT April 10 sampling, we consistently find that the diameter is 4 μas smaller than the original image. The origin of this discrepancy is discussed in Figure 7. For the simulated data using the April 10 sampling, which has poor coverage compared with the other days, the imaging gives two modes. One mode is similar to the other days, while the second fails to show a coherent disk structure giving the second peak in diameter at 60 μas.

The ability of VIDA has been demonstrated for several sources, each with over a thousand reconstructions. These include image reconstructions from simulated data produced from double Gaussians, slashed rings, and GRMHD simulations. In all cases, key quantitative features were accurately recovered where they appeared in the underlying image reconstructions. These include separations, orientations, ring diameters, widths, brightness profiles, and multiple measures of ellipticity. Application of these to the EHT observations of M 87 will be explored in future work.

The applicability of VIDA extends beyond EHT observations of M 87. The ability to create composite models with multiple components is naturally relevant to VLBI reconstruction of an active galactic nucleus, such as 3C 279, that is composed of multiple compact features (e.g., Kim et al. 2020).

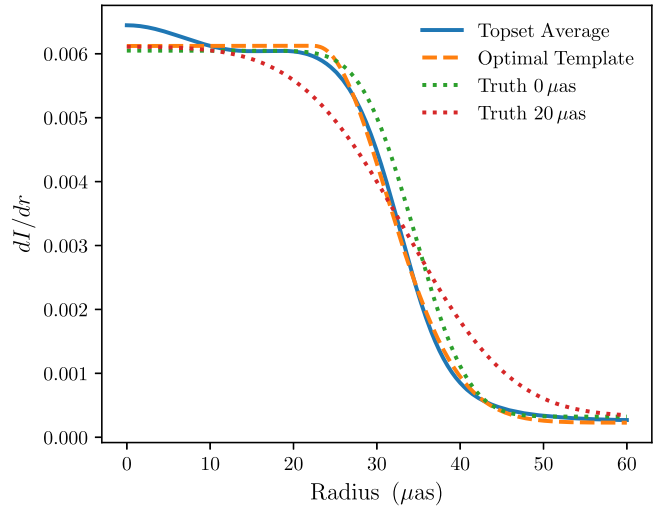

Figure 7. Results when applying VIDA to the average reconstruction from the top set using the simulated April 11 EHT sampling. The images were normalized to unit flux and centered before averaging. The average radial profile is shown in solid blue. Comparing this to the optimal template (orange dashed line) and the true profile (dotted lines), we see that the optimal template matches the reconstruction's radial profile but underestimates the ground-truth image.

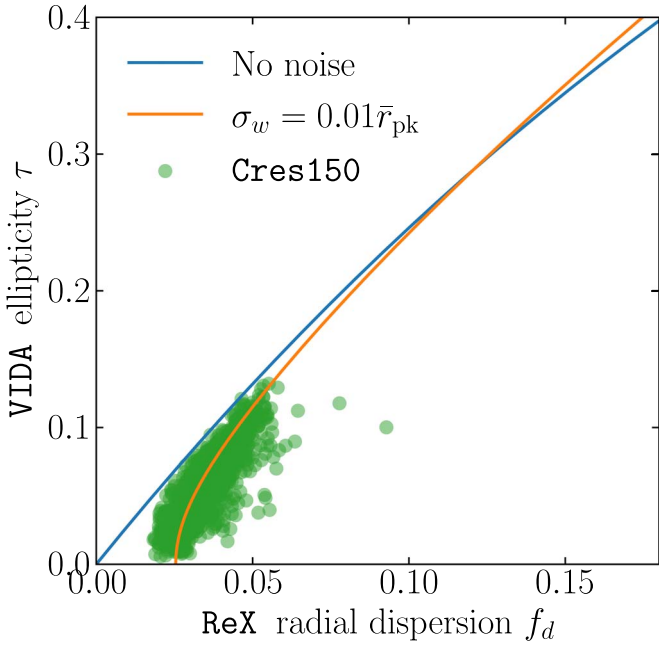

Figure 8. Comparison of the REx's fractional radial dispersion and VIDA's τ ellipticity parameter. The blue curve shows the conversion for the case of an ellipse, the orange curve is an ellipse with a constant white noise fluctuation in the found radius, with a variance of $0.01^2\bar{r}_{\rm pk}^2$, and the green points are the results of fitting the Cres150 top set in Section 3 with both REx and VIDA.

It should be noted that image feature-extraction methods, like VIDA, are generally most useful when strong priors may be placed on the image structure itself. That is, VIDA is primarily a method for quantifying what is already qualitatively apparent. Poorly chosen models can lead to significant parameter biases, as seen in Section 4.1, where an extra Gaussian blob was required to achieve acceptable results. However, because VIDA is an image characterization tool, not an imaging tool in and of itself, this presents only a very modest limitation on its utility.

We thank George Wong, Ben Prather, and Charles Gammie for kindly providing synthetic images based on GRMHD simulations. We would also like to thank the referees for their helpful comments. This work was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Economic Development, Job Creation and Trade. P.T. receives support from the Natural Science and Engineering Research Council through the Alexander Graham Bell CGS-D scholarship. A.E.B. thanks the Delaney Family for their generous financial support via the Delaney Family John A. Wheeler Chair at Perimeter Institute. A.E.B. receives additional financial support from the Natural Sciences and Engineering Research Council of Canada through a Discovery Grant. We thank the National Science Foundation (grant Nos. AST-1716536, AST-1440254, AST-1935980) and the Gordon and Betty Moore Foundation (grant No. GBMF-5278) for financial support of this work. This work was supported in part by the Black Hole Initiative, which is funded by grants from the John Templeton Foundation and the Gordon and Betty Moore Foundation to Harvard University.

Software: BlackBoxOptim.jl, eht-imaging, GR (Heinen 2022), Julia (Bezanson et al. 2017), matplotlib 3.3 (Hunter 2007), Pandas (McKinney 2010; Reback et al. 2021), Python 3.8.3 (Van Rossum & Drake 2009), Scipy (Virtanen et al. 2020), ThemisPy, VIDA.jl.

Appendix A
REx Ring Parameter Definitions

The first step in REx (see Paper IV, for details) is to identify the dominant ring in the image. Given a center position (x, y), REx samples the image in radius r and azimuthal angle θ obtaining an intensity map $I(r, \theta|x, y)$. Then for that central map, the radius of the ring is taken as the azimuthally averaged peak brightness:

$$r_{\rm pk}(\theta|x, y) = \mathrm{argmax}[I(r, \theta|x, y)]_r$$
$$\bar{r}_{\rm pk} = \langle r_{\rm pk}(\theta|x, y)\rangle_{\theta\in[0,2\pi]}. \quad \text{(A1)}$$

This provides a different "radius" for every point (x, y) in the image. To find the dominant ring in the image, the fractional radius spread is minimized:

$$(x_0,\ y_0) = \mathrm{argmin}\left[\frac{\sigma_{\bar{r}(x,y)}}{\bar{r}_{\rm pk}(x, y)}\right], \quad \text{(A2)}$$

where $\sigma_{\bar{r}}(x, y) = \langle(r_{\rm pk}(\theta|x, y) - \bar{r}_{\rm pk})^2\rangle$ is the radial dispersion. The diameter in the image is then given by

$$d = 2\bar{r}_{\rm pk}(x_0, y_0). \quad \text{(A3)}$$

To relate this to VIDA's definition we consider an ellipse with semimajor axis a and semiminor axis b. Then VIDA parameterizes this ellipse with $d_0 = 2r_0 = 2\sqrt{ab}$ and $\tau = 1 - b/a$. The relationship between r_0 and $r_{\rm pk}$ is then given by

$$\bar{r}_{\rm pk} = \frac{r_0}{\sqrt{1-\tau}}\frac{1}{2\pi}\int_0^{2\pi}\sqrt{1-\epsilon^2(\tau)\sin^2(\theta)}\,{\rm d}\theta = \frac{2}{\pi}\frac{r_0}{\sqrt{1-\tau}}E(\epsilon(\tau)), \tag{A4}$$

where $E(x)$ is the complete elliptic integral of the second kind and $\epsilon(\tau) = \sqrt{1-(1-\tau)^2}$ is the orbital eccentricity.

REx's measure of circularity of the ring is provided by the radial fractional dispersion:

$$f_d = \frac{\sigma_{\bar{r}}}{\bar{r}_{\rm pk}}. \tag{A5}$$

To compare REx's ellipticity measure, f_d, to VIDA's, we need to relate f_d to $\tau = 1 - b/a$:

$$f_d(\tau) = \frac{\sqrt{1-\epsilon(\tau)^2 - 4/\pi^2 E^2(\epsilon)}}{\sqrt{1-\tau}} = \frac{\sqrt{(1-\tau)^2 - 4/\pi^2 E^2(\epsilon)}}{\sqrt{1-\tau}}. \tag{A6}$$

Equation (A5) is then used to convert this to a fractional diameter spread. Using linear interpolation, we invert the function achieving a map from f_d to τ. One important thing to note is that this conversion assumes that the image is a perfect ellipse. In general, this will not be true for the image reconstructions, and we expect that the REx ellipticity may be different than the VIDA measurement. To model this, we consider $r_{\rm pk}$ modified by a white noise ϵ_θ term with dispersion proportional to the average radius $\langle \epsilon_\theta \epsilon_{\theta'}\rangle = \sigma_w^2 \bar{r}_{\rm pk}^2 \delta(\theta - \theta')$. In this base the average peak radius is unchanged since ϵ has mean 0. However, the additional noise does impact the radial dispersion:

$$\sigma_{\bar{r}} \to \sigma_{\bar{r}} + 2\langle \epsilon_\theta r_{\rm pk}(\theta)\rangle + \langle \epsilon_\theta^2\rangle. \tag{A7}$$

When we have a circular ring then this just becomes $\langle \epsilon_\theta^2\rangle = \sigma_\epsilon^2 \bar{r}_{\rm pk}^2$ adding a constant floor. Figure 8 shows the conversion when the ring is elliptical and compares it to the results of the Cres150 top set of Section 3.

The width of the ring is defined by finding the FWHM at a fixed θ ray and then averaging over θ:

$$w = \langle {\rm FWHM}_r[I(r,\,\theta|x_0,\,y_0) - I_{\rm floor}]\rangle_\theta. \tag{A8}$$

The intensity floor is given by $I_{\rm floor} = \langle I(r = 50\mu{\rm as},\,\theta)\rangle_\theta$ and is included to avoid biasing the measurement due to the low-level emission present in the image. This is similar to including the constant intensity template during the VIDA extraction.

In order to characterize the azimuthal profile of the ring (ξ_s and s for VIDA), we consider the azimuthal moments of the ring. Namely, the orientation ξ_s is given by

$$\xi_s = \left\langle {\rm Arg}\left[\int_0^{2\pi} I(r,\,\theta|x_0,\,y_0)e^{i\theta}{\rm d}\theta\right]\right\rangle_{r\in[r_{\rm in},r_{\rm out}]}, \tag{A9}$$

where $r_{\rm in} = (d-w)/2$ and $r_{\rm out} = (d+w)/2$. The strength of the slash is given by

$$s = 2\left\langle \frac{\left|\int_0^{2\pi} I(r,\,\theta|x_0,\,y_0)e^{i\theta}{\rm d}\theta\right|}{\int_0^{2\pi} I(r,\,\theta|x_0,\,y_0){\rm d}\theta}\right\rangle. \tag{A10}$$

Note that the factor of 2 is included to match VIDA's definition (Equation (4)).

Appendix B
Joint Density Plot for Crescent 150 Reconstruction

The joint density plot from the top set image ensemble are expected to be correlated. Figure 9 shows the joint density plot for the crescent 150 image reconstructions for both the VIDA and REx pipelines.

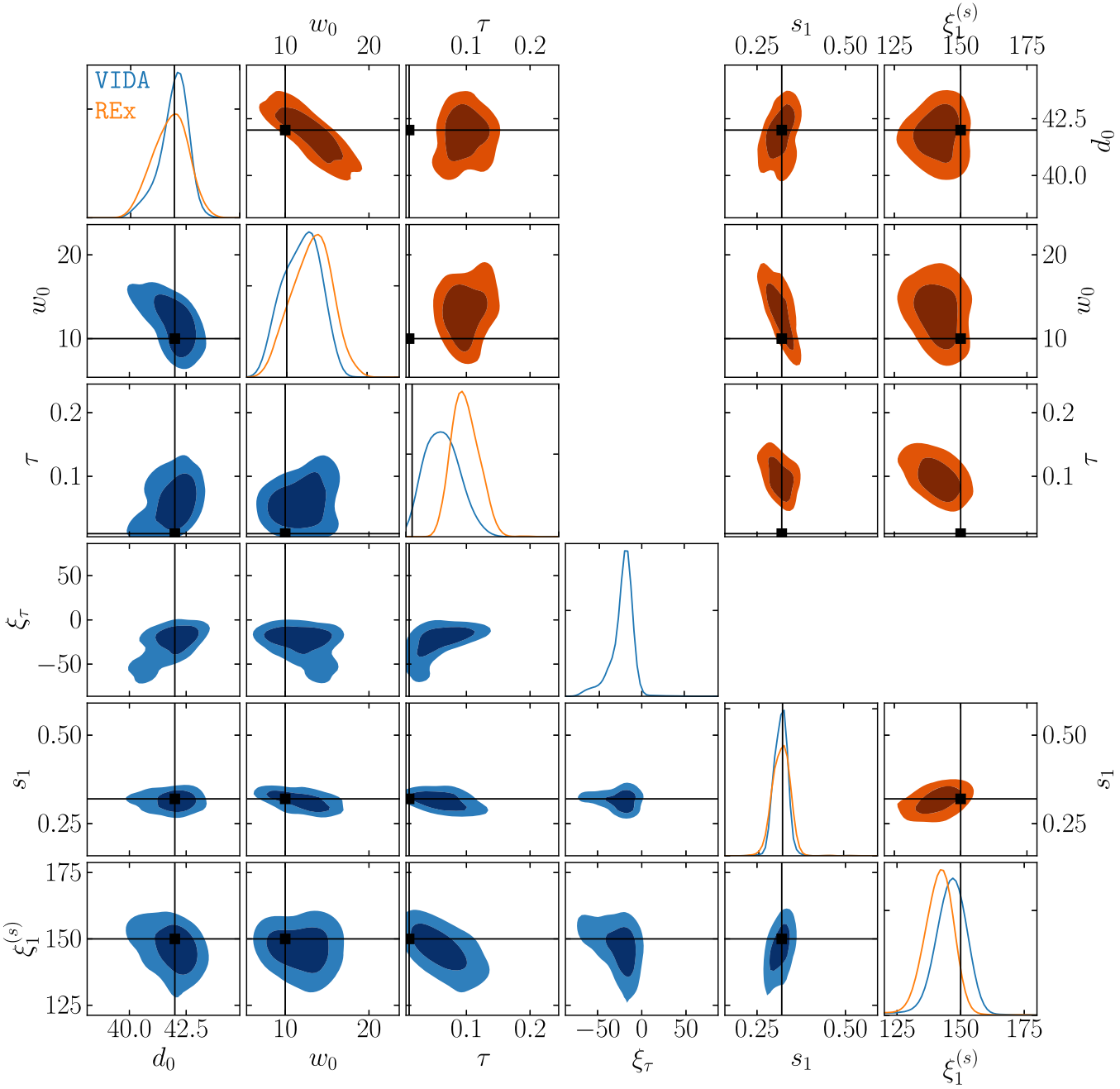

Figure 9. Joint density plot for the crescent image reconstruction ensemble using the April 11 EHT simulated coverage. For VIDA (blue, lower), we find that the truth values (black lines) are within the recovered distributions. For REx (orange, upper) we find that for all parameters except the ellipticity τ are recovered. REx's ellipticity has a floor around $\tau = 0.05$, which is expected since REx measures fractional dispersion as is explained in Appendix A. Additionally, no ξ_τ is measured for REx, so those columns are not shown.

Appendix C
CLEAN Reconstruction Results

In this paper we have focused on RML methods. The reason for this twofold. First, the RML results consisted of the majority of the methods used in Paper IV. Second, and more important, RML methods produced thousands of simulated image reconstructions. The DIFMAP/CLEAN pipelines used in Paper IV only produce 30 images per simulated image. This effectively increases the sampling uncertainty in the reconstructions. Nevertheless, in this section we will apply VIDA and REx to the CLEAN reconstructions of the crescent with 180° position angle and the GRMHD simulation from Section 3.4. Afterward we will repeat the analysis of the disk and double simulated EHT data from Section 4 solely using VIDA. We will only present results on the EHT sampling of April 11, but we found similar results on the other days.

C.1. Ring-like Reconstructions

To analyze the CLEAN reconstructions we followed an identical procedure as in Section 3, except we used the CLEAN image reconstructions pipeline. An example reconstruction and optimal VIDA template for the crescent and GRMHD simulation are shown in the first two columns of Figure 10. We see that the crescent is accurately recovered, although it is generally thicker than the on-sky image, which is due to the CLEAN beam.

Quantitatively, the marginal parameter distributions for the crescent are shown in the top panels of Figure 11. From these

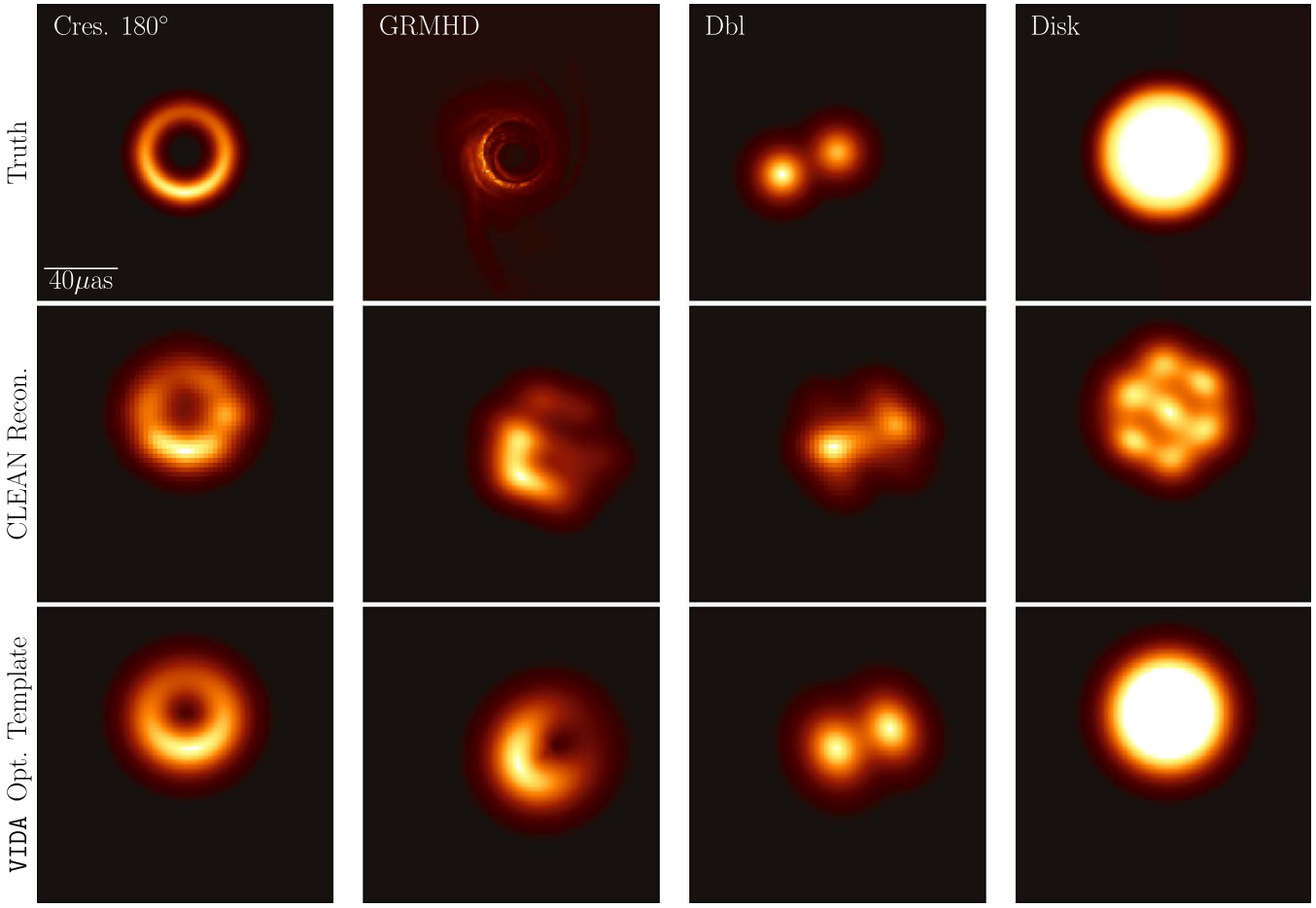

Figure 10. Example reconstructions and optimal VIDA templates when applied to the crescent 180° (first column), GRMHD simulations (second column), double (third column), and disk (fourth column). The top row shows the on-sky image; the middle row, an example CLEAN reconstruction; and the bottom row, the optimal VIDA template.

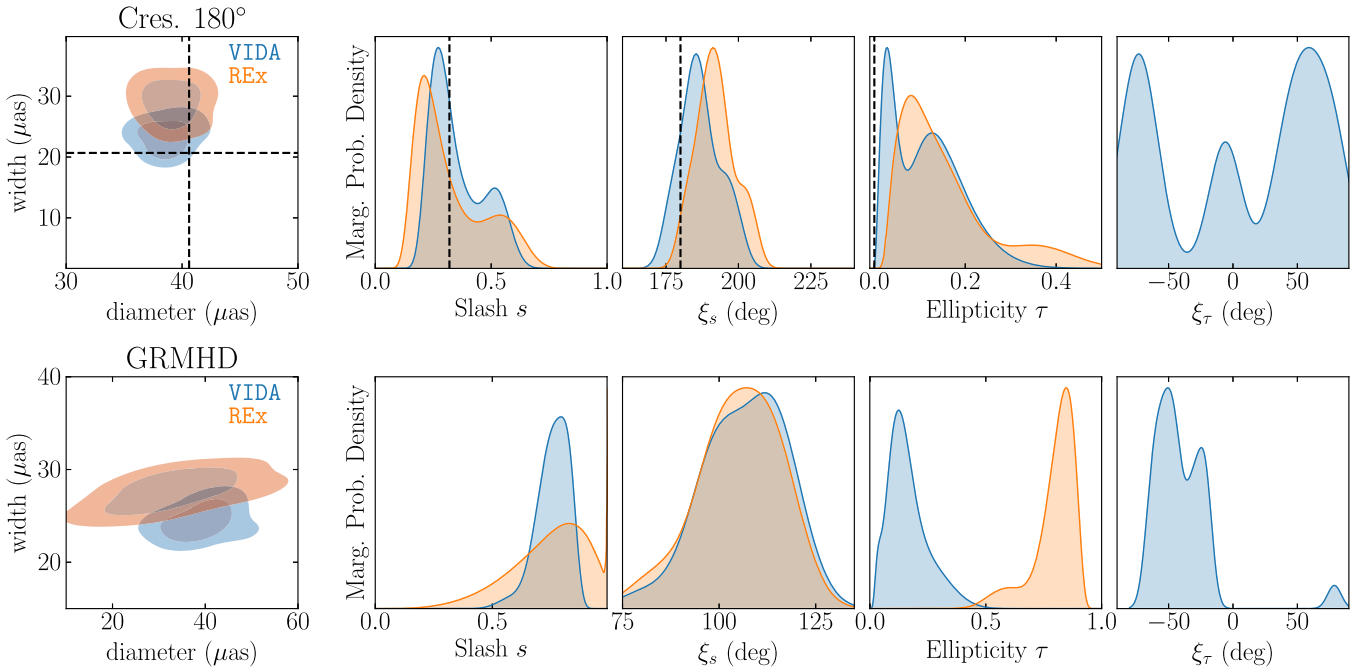

Figure 11. Marginal distributions from applying VIDA (blue) and REx (orange) to the CLEAN reconstructions. The top row shows the crescent with brightness position angle 180° east of north. Note we see similar results for the other crescent orientation. The bottom row shows the results from the GRMHD simulations. Here we see large differences in the measured parameters between REx and VIDA. The origin of these differences is likely from the GRMHD simulations lacking a dominant ring in the image, breaking the underlying ring assumptions of REx and VIDA.

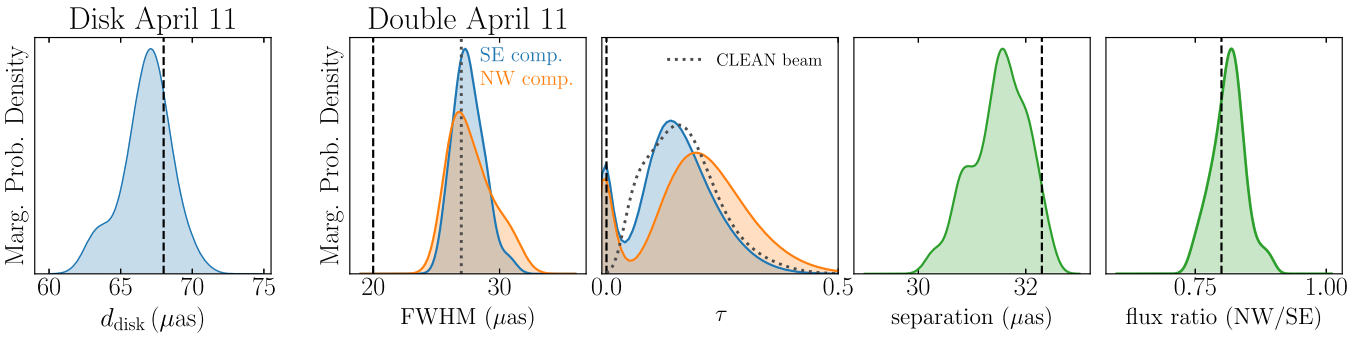

Figure 12. Marginal distribution from applying VIDA to the disk (left panel) and double (remaining panels) simulated data sets. In this case we see that CLEAN is able to recover the diameter of the disk unlike eht-imaging, which had underestimated the diameter by a few microarcseconds (see Figure 10). For the double we see that VIDA and the CLEAN reconstructions are able to recover the separation and flux ratio. The true ellipticity is also recovered. Note that the larger τ and FWHM for the Gaussians follow from incorporating the contributions of the CLEAN beam (gray dotted line).

reconstructions we see a number of interesting differences when compared with the eht-imaging results in Figure 3. The recovered crescent diameter is generically smaller than the eht-imaging results. This difference can be explained by the larger width. That is, the measured on-sky width is given by the rms of the CLEAN beam size and the intrinsic ring width, which is shown by the black lines in Figure 11. The measured diameter then matches the diameter of the peak brightness after blurring, that is, Equation (11). One interesting thing to note is that the ellipticity for the CLEAN reconstructions appears not to have the same north–south orientation bias as the eht-imaging reconstructions. This suggests that the ellipticity bias we saw for the eht-imaging reconstructions is due to the specific hyperparameters used in the construction of the top set. Note that while we did not show the crescent with position angle 150°, we did get similar results.

For the GRMHD results shown in the bottom row of Figure 11, the distributions between REx and VIDA are in disagreement. Namely, the REx diameter distribution width is quite a bit larger. Additionally, the measured ellipticity is discrepant between the two feature-extraction methods. This occurs because the CLEAN reconstructions of the GRMHD simulation do not show a dominant ring feature (see Figure 10). In this case, the assumptions underlying REx's parameter definitions and VIDA's ring template break down. We do note that qualitatively VIDA does seem to recover the visual image size and ellipticity better than REx. Furthermore, in this case a user would probably employ a different VIDA template given the apparent nonring structure, especially since the actual on-sky image structure would not be known a priori.

C.2. Disk and Double Reconstructions

For the disk we used the same template as in Section 4.2. However, for the double we only used a two-component asymmetric Gaussian template. The reason for this was that the CLEAN reconstructions only had two Gaussian components, unlike the eht-imaging ones. We show example reconstructions and the optimal template for the disk and double in Figure 10. Visually we see that the templates are reconstructing the qualitative features of both the double and disk. More quantitatively, we show the marginal parameter distributions in Figure 12. We find near-identical FWHMs for both Gaussian components, although both are larger than the ground truth, due to the larger CLEAN beam used. However, the rest of the parameters are consistent with the truth in all cases. For the disk, we find that unlike the eht-imaging reconstructions, the disk size is included in the marginal distribution. Namely, the CLEAN reconstructions do not seem to suffer from the same size bias as the eht-imaging top set. This further supports our hypothesis that the size bias is due to the eht-imaging hyperparameters chosen in Paper IV.

ORCID iDs

Paul Tiede https://orcid.org/0000-0003-3826-5648
Avery E. Broderick https://orcid.org/0000-0002-3351-760X
Daniel C. M. Palumbo https://orcid.org/0000-0002-7179-3816

References

Aherne, F. J., Thacker, N. A., & Rockett, P. I. 1998, Kybernetika, 34, 363
Akiyama, K., Kuramochi, K., Ikeda, S., et al. 2017a, ApJ, 838, 1
Akiyama, K., Ikeda, S., Pleau, M., et al. 2017b, AJ, 153, 159
Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. 2017, SIAMR, 59, 65
Bhattacharyya, A. 1943, Bull. Calcutta Math. Soc., 35, 99
Blei, D. M., Kucukelbir, A., & McAuliffe, J. D. 2017, Journal of the American Statistical Association, 112, 859
Bouman, K. L., Johnson, M. D., Zoran, D., et al. 2016, in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), ed. M. Herbordt, M. Leeser, & L. Shannon (Piscataway, NJ: IEEE)
Broderick, A. E., Pesce, D. W., Tiede, P., Pu, H.-Y., & Gold, R. 2020, ApJ, 898, 9
Chael, A. A. 2019, PhD thesis, Harvard University, Graduate School of Arts & Sciences
Chael, A. A., Johnson, M. D., Bouman, K. L., et al. 2018, ApJ, 857, 23
Chael, A. A., Johnson, M. D., Narayan, R., et al. 2016, ApJ, 829, 11
Choi, E., & Lee, C. 2003, PatRe, 36, 1703
Clark, B. G. 1980, A&A, 89, 377
Das, S., & Suganthan, P. N. 2011, IEEE Transactions on Evolutionary Computation, 15, 4
Duda, R. O., & Hart, P. E. 1972, Commun. ACM, 15, 11
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019a, ApJL, 875, L1
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019b, ApJL, 875, L2
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019c, ApJL, 875, L3
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019d, ApJL, 875, L4
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019e, ApJL, 875, L6
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019f, ApJL, 875, L5
Frieden, B. R. 1972, JOSA, 62, 511
Goffe, W. L. 1996, Studies in Nonlinear Dynamics & Econometrics, 1, 1
Goudail, F., Réfrégier, P., & Delyon, G. 2004, JOSAA, 21, 1231
Gull, S. F., & Daniell, G. J. 1978, Natur, 272, 686

Heinen, J., et al. 2022, GR Framework v0.63.1, https://github.com/sciapp/gr
Högbom, J. A. 1974, A&AS, 15, 417
Honma, M., Akiyama, K., Uemura, M., & Ikeda, S. 2014, PASJ, 66, 95
Hough, P. V. C. 1964, OptEn, 2, 379
Hunter, J. D. 2007, CSE, 9, 90
Ikeda, S., Tazaki, F., Akiyama, K., Hada, K., & Honma, M. 2016, PASJ, 68, 45
Kim, J., Krichbaum, T. P., Broderick, A. E., et al. 2020, A&A, 640, A69
Kullback, S., & Leibler, R. A. 1951, Ann. Math. Statist., 22, 79
Kuramochi, K., Akiyama, K., Ikeda, S., et al. 2018, ApJ, 858, 56
McKinney, W. 2010, in Proc. 9th Python in Science Conf. 445, ed. S. van der Walt & J. Millman (Austin, TX: SciPy), 56
Narayan, R., & Nityananda, R. 1986, ARA&A, 24, 127
Reback, J., Mendel, J. B., & McKinney, W. 2021, Pandas v.1.3.0, Zenodo, doi:10.5281/zenodo.3509134
Schwab, F. R. 1984, AJ, 89, 1076
Schwarz, U. J. 1978, A&A, 65, 345
Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 2017, Interferometry and Synthesis in Radio Astronomy (3rd ed.; Berlin: Springer)
Tiede, P., Broderick, A., Palumbo, D., & Chael, A. 2020, arXiv:2012.07889
Van Rossum, G., & Drake, F. L. 2009, Python 3 Reference Manual (Scotts Valley, CA: CreateSpace)
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261