

1 Generative AI for image reconstruction in Intensity Interferometry: a first attempt

2 KM NITU RAI,^{1,2} YURI VAN DER BURG,³ SOUMEN BASAK,¹ PRASENJIT SAHA,³ AND SUBRATA SARANGI^{4,5}

3 ¹*School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram, Maruthamala PO, Vithura,*
4 *Thiruvananthapuram 695551, Kerala, India.*

5 ²*Aryabhatta Research Institute of Observational Sciences, Manora Peak, Nainital 263129, India.*

6 ³*Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.*

7 ⁴*School of Applied Sciences, Centurion University of Technology and Management, Odisha-752050, India.*

8 ⁵*Visiting Associate, Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, Maharashtra,*
9 *India.*

ABSTRACT

In the last few years Intensity Interferometry (II) has made significant strides in achieving high-precision resolution of stellar objects at optical wavelengths. Despite these advancements, phase retrieval remains a major challenge due to the nature of photon correlation. This paper explores the application of a conditional Generative Adversarial Network (cGAN) to tackle the problem of image reconstruction in II. This method successfully reconstructs the shape, size, and brightness distribution of simulated, fast-rotating stars based on a sparsely sampled spatial power spectrum obtained from a hypothetical ground-based II facility composed of four Imaging Atmospheric Cherenkov Telescopes (IACTs). Although this particular example could also be addressed using parameter fitting, the results suggest that with larger arrays much more complicated systems could be reconstructed by applying machine-learning techniques to II.

20 1. INTRODUCTION

21 Humans instinctively feel a relationship with the stars.
22 One of the primary scientific projects of humanity is to
23 figure out what the stars are and how do they do what
24 they do. The first obvious step in this project, beyond
25 measuring their global parameters like diameter, mass,
26 orbital and astrometric elements, is to obtain images of
27 the stars with all details of the stellar surfaces. In case
28 of the Sun, this is done routinely by observatories and
29 sun-observation satellites. But such routine still remains
30 a challenge even for the α -centauri system, our nearest
31 stellar neighbour. The objective is to achieve capability
32 of high fidelity image reconstruction of distant stars.
33 Two interferometry based techniques, namely, Michelson
34 Interferometry (MI) or Intensity Interferometry (II)
35 have emerged during the last century to address this
36 objective. A discussion of the development of these two
37 approaches and comparison of their respective merits
38 and challenges can be found in (K. N. Rai et al. 2025).
39 The work reported here presents the results of a first
40 effort at applying a Conditional Generative Adversarial
41 (neural) Network (cGAN) to image reconstruction of a
42 fast rotator using its simulated II observations.

43 The foundational basis of II stems from the pioneering
44 experiments and theoretical investigations, widely
45 referred to as the “HBT effect”, carried out initially by

46 Hanbury Brown and Twiss (HBT) R. Hanbury Brown
47 & R. Q. Twiss (1956); R. Hanbury Brown et al. (1957,
48 1958) and, later, by Glauber (R. J. Glauber 1963). The
49 correlation between photons (called HBT effect) measured
50 by a pair of photon detectors in two coherent
51 beams of light was reported in 1956 R. Hanbury Brown
52 & R. Q. Twiss (1956). K. N. Rai et al. (2025) present
53 a modern-day variant of this experiment, carried out
54 with pseudo-thermal light. The HBT Effect and the re-
55 lated theoretical investigations laid the foundation for
56 the modern field of Quantum Optics.

57 Hanbury Brown and his collaborators led the creation
58 and installation of the iconic II facility at Narrabri, Aus-
59 tralia and reported the measurement of angular diame-
60 ters of 32 stars and a few multiple star-systems (R. Han-
61 bury Brown et al. 1974). Soon after this work, however,
62 II observations of stars was stalled for over four decades
63 due to the limits of the then available photon detectors
64 and data processing equipment. With gradual mitiga-
65 tion of such issues, proposals to utilize Imaging Atmo-
66 spheric Cherenkov Telescope (IACT) facilities for con-
67 ducting II observations of stars have emerged (J. Le Bo-
68 hec 2006; P. D. Nuñez et al. 2010, 2012a; D. Dravins et
69 al. 2013) as a secondary science application of these fa-
70 cilities during moonlit nights. SII observations at VER-
71 ITAS, MAGIC, and HESS are now being reported (e.g.,

A. Acharyya et al. 2024; S. Abe et al. 2024; N. Vogel et al. 2025). This approach has the potential to enhance the scientific output of existing IACT facilities, and especially of the upcoming Cherenkov Telescope Array Observatory (CTAO). Simulations carried out by Rai et al. (K. N. Rai et al. 2021, 2022, e.g.,) have shown that recent advancements in photon detectors could be effective in achieving high-precision measurements of parameters for stellar objects.

The thrust of these efforts has been to measure the average global parameters of stars and star systems. The high fidelity imagery, however, would transcend the measurement of global and average stellar parameters, such as angular diameters, binary separations, and orbital characteristics, which offer only an integrated view of the star or star system as a whole. Such imaging capabilities promise direct insights into the dynamic surface phenomena, including limb darkening, convection cells, granulation, star spots, oblateness and gravity darkening in rapid rotators and atmospheric structures, akin to the detailed observations routinely conducted on our own Sun.

As it stands today, studies grappling various issues of image reconstruction are being reported (X. Haubois et al. 2009; R. P. Norris et al. 2021; L.-C. Liu et al. 2024; L.-C. Liu et al. 2025). MI-based image reconstruction has made substantial progress in this area with efforts at generating constructed images of stars like Betelgeuse (X. Haubois et al. 2009) and AZ Cyg (R. P. Norris et al. 2021). On the other hand, II-based methods are in a nascent stage. The recent publications L.-C. Liu et al. (2024); L.-C. Liu et al. (2025) have demonstrated, through outdoor experiments, imaging millimeter-scale targets at 1.36 km with a resolution 14 times better than a single telescope's diffraction limit. A "flexible computational algorithm" reconstructs images from intensity correlations, overcoming atmospheric turbulence and optical imperfections.

We report here, the results of our attempt – the first of its kind – to construct the gravity-darkened sky-image of fast rotating stars consistent with their respective simulated ground-based II-observations using a cGAN neural network architecture. Image reconstruction in gravity darkened fast rotating stars has long been examined using various methods in MI (G. T. van Belle et al. 2001; A. Domiciano de Souza et al. 2003, 2005; H. A. McAlister et al. 2005; J. D. Monnier et al. 2007; E. Pedretti et al. 2009; M. Zhao et al. 2009; A. O. Martinez et al. 2021). Recently photosphere oblateness of γ -Cassiopeia (A. Archer et al. 2025) has been measured at the VERITAS observatory using II. These results, and especially, that by A. Archer et al. (2025) put our work in con-

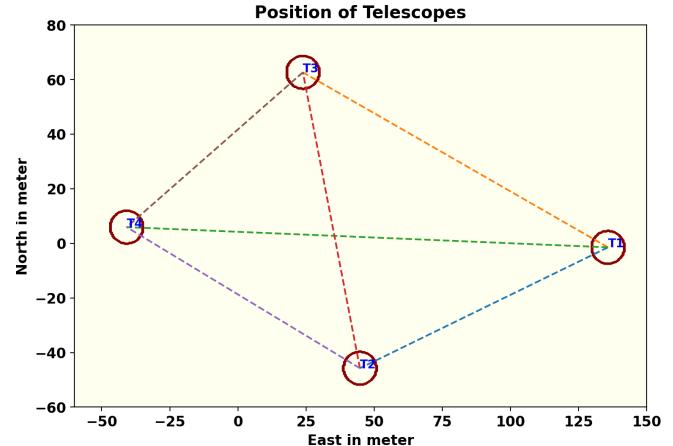


Figure 1. A schematic representation of a hypothetical observation facility with an array of four Cherenkov Telescopes (IACTs) whose relative positions approximately mimic those at VERITAS. This array is used in simulating the II observation of the fast rotators. Each of the telescopes has a diameter of 12m. The baselines provided by the array are of the order of 100m.

text. Thus our work presented here is a natural next step of the work of Archer et al and others. We emulate the cGAN model P. Isola et al. (2017) to reconstruct images of fast-rotating stars using their simulated Intensity Interferograms and sky-intensity distributions as input data for training, testing, and validation. We consider an array of four Imaging Cherenkov Telescopes (IACTs) whose relative positions approximately mimic those at VERITAS and simulate observations of a set of synthetically generated fast-rotating stars. The image predicted by the trained cGAN shows promising results in reconstructing the stars' shapes and sizes. The reconstructed brightness distributions are then assessed using moments.

This paper is organized as follows. The next section discusses briefly the past efforts at image reconstruction on II, followed by the section on a discussion of II, focusing on its signal and noise characteristics for fast-rotating stars along the Earth's rotation. The following section introduces the cGAN formulation and its structure. The fifth section details the parameter selection for training the network for image reconstruction. The sixth section presents the results of the trained network both visually and via image moments. Finally, the paper concludes with a discussion of the overall results.

151 2. PAST EFFORTS AT IMAGE
 152 RECONSTRUCTION IN INTENSITY
 153 INTERFEROMETRY

154 Several approaches have been developed for phase re-
 155 construction in intensity interferometry (II). H. Gamo
 156 (1963) introduced triple-intensity correlation, applied by
 157 M. L. Goldberger et al. (1963) to microscopic systems
 158 and extended by T. Sato et al. (1978, 1979, 1981) to
 159 measure stellar diameters and phases, though limited
 160 by low signal-to-noise ratio (SNR). R. W. Gerchberg
 161 (1972) proposed an iterative phase retrieval method us-
 162 ing image and diffraction plane data, sensitive to initial
 163 estimates and convergence speed. J. Fienup (1982) im-
 164 proved this with the Hybrid Input-Output algorithm,
 165 enhancing robustness in noisy conditions. R. Holmes et
 166 al. (2010) utilized Cauchy-Riemann relations for 1-D
 167 and 2-D image reconstruction, applied to simulated
 168 stellar data with Imaging Cherenkov Telescope Arrays
 169 (P. D. Nuñez et al. 2010, 2012a,b), but faced compu-
 170 tational complexity for higher dimensions. X. Li et al.
 171 (2014) introduced a regularized iterative method incor-
 172 porating priors (e.g., sparsity) to mitigate noise and
 173 ill-posedness, though challenged by parameter tuning
 174 and initial guess sensitivity. The Transport-of-Intensity
 175 Equation (TIE), proposed by M. R. Teague (1983), re-
 176 trievals phase from intensity variations across planes;
 177 J. Zhang et al. (2020) solved TIE as a Poisson equation
 178 using a maximum intensity assumption, while C.
 179 Kirisits et al. (2024) combined TIE with the Trans-
 180 port of Phase Equation for improved accuracy across
 181 arbitrary apertures and non-uniform illumination, with
 182 convergence dependent on initial guesses and boundary
 183 conditions.

184 With non-linearity built into their architecture, artifi-
 185 cial neural networks (ANNs) empowered by deep learn-
 186 ing methods are promising for exploring the task of re-
 187 constructing images of stellar objects from ground-based
 188 observations. Convolutional Neural Networks (CNNs),
 189 with their specialized architecture for processing two-
 190 dimensional datasets, are a natural choice for image
 191 processing tasks. In astronomical image reconstruction
 192 projects, a common challenge is that the interferometric
 193 data are typically undersampled as well as noisy. There-
 194 fore, the CNN architectures and deep learning methods
 195 employed must be capable of reliably learning both the
 196 global context of the training dataset and the local fea-
 197 tures within it. Among the various CNN architectures,
 198 U-Net models (O. Ronneberger et al. 2015) have proven
 199 successful in such tasks.

200 Furthermore, given that achieving a high signal-to-
 201 noise ratio (SNR) is often challenging in astronomical
 202 datasets, it is immensely beneficial if additional data can

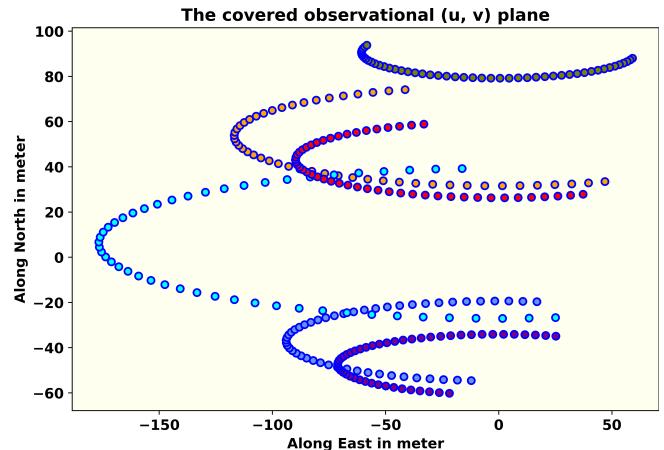


Figure 2. The tracks of the baselines due to Earth rotation (described in sec. 3.3) provided by the hypothetical observation facility of four telescopes arranged in fig. 1 for one night of observation. The number of baselines scales as square of number of telescopes in the array thus leading to greater coverage of the observational plane and better image reconstruction prospects.

203 be generated using the available information from the
 204 observed sky density distribution and ground-based ob-
 205 servations (II data, in our case) of the sources under in-
 206 vestigation. Generative Adversarial Networks (GANs),
 207 introduced by I. Goodfellow et al. (2014), have been
 208 successful in such data augmentation tasks. Conditional
 209 GAN (cGAN) architectures, proposed by M. Mirza et al.
 210 (2014) and applied to a wide variety of datasets by P.
 211 Isola et al. (2017), leverage additional information about
 212 the images in the training datasets and have demon-
 213 strated remarkable robustness in image recovery across
 214 diverse data types.

215 In the astrophysical context, K. Schawinski et al.
 216 (2017) employed a GAN model to recover features, such
 217 as spiral arms, central bulges, and disk structures of
 218 galaxies, from noise-affected images. M. Mustafa et
 219 al. (2019) developed and customized a Deep Convo-
 220 lutional GAN, dubbed ‘‘CosmoGAN’’, capable of gen-
 221 erating high-fidelity weak-lensing convergence maps of
 222 dark matter distribution that statistically reproduce real
 223 weak lensing structures. D. Coccolini et al. (2021)
 224 have successfully generated credible images of planets,
 225 nebulae, and galaxies using ‘‘lightweight’’ and ‘‘physics-
 226 uninformed’’ GANs to produce synthetic images of celes-
 227 tial bodies. They also generated a ‘‘Hubble Deep Field-
 228 inspired’’ wide-view simulation of the universe.

229 3. INTENSITY INTERFEROMETRY (II) WITH
 230 IACT ARRAYS

231 This section presents a brief conceptual overview of
 232 how an array of telescopes is used to perform II obser-

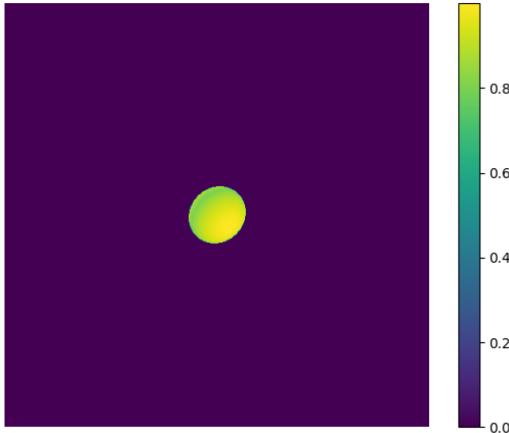


Figure 3. This figure shows one of the simulated fast rotating stars (FRS). The brightness is highest at the poles; gravity darkening visible along the equator. A total of 31460 such images of FRS with different parameter values have been generated to train the model.

vations, and explains the Signal-to-Noise Ratio (SNR) from these measurements.

3.1. The signal for II

As a simple example, let us consider a pair of IACTs pointed at a star. Suppose the two telescopes simultaneously measure the intensity of radiation $I_1(t)$ and $I_2(t)$, respectively. The signals from these detectors are cross-correlated and averaged over time, yielding the second order ($n = 2$) correlation of these intensities as (cf. V. A. Acciari et al. 2020; D. Dravins et al. 2013)

$$g^{(2)} = \frac{\langle I_1(t) \cdot I_2(t + \tau) \rangle}{\langle I_1(t) \rangle \cdot \langle I_2(t) \rangle} \quad (1)$$

where τ is the time delay between the telescopes. For spatially coherent and randomly polarized light, Eq. (1) reduces to the relation (sometimes called the Siegert relation, see e.g., V. A. Acciari et al. 2020).

$$g^{(2)} = 1 + \frac{\Delta f}{\Delta \nu} |V_{12}|^2 \quad (2)$$

where Δf is the electronic bandwidth of the photon detectors which measure the intensities and $\Delta \nu$ is the frequency bandwidth of the filters employed in the telescopes to observe the star. Values of $\Delta \nu \sim 1$ THz and $\Delta f \sim 1$ GHz are typical of recent work. In Eq. (2), V_{12} , referred to as the complex visibility function, is the Fourier transform of the source brightness distribution. It contains information about the star's angular diameter. However, the phase information is lost since we measure only the absolute value $|V_{12}|^2$. In observational

astronomy, the correlation is often expressed in terms of the normalized contrast, given by:

$$c = g^{(2)} - 1 = \frac{\Delta f}{\Delta \nu} |V_{12}|^2 \quad (3)$$

with $|V_{12}|^2$ being a function of baseline $b = \sqrt{u^2 + v^2}$ on the observational plane (u, v). This implies the strength of the signal would be enhanced if a larger number of baselines or pairs of telescopes are employed.

3.2. The Signal-to-Noise Ratio for II

The primary purpose of IACTs is to study high-energy gamma rays (with energy $E \geq 30$ GeV) arriving from cosmic sources, entering the Earth's atmosphere, and initiating Cherenkov showers in the upper atmosphere due to multiple scattering. These telescopes feature an array of mirrors that focus light received from a sky source onto their respective set of photo-multiplier tubes (PMTs, see e.g., J. Aleksić et al. 2016) with appropriate specifications needed for II observations. In the simulation model adopted here, we consider a set of four IACTs, each with similar properties. The positional configuration of these IACTs is shown in Fig. 1. The optical signal directed to a PMT is filtered using a spectral filter with a chosen mean observational wavelength λ and corresponding bandpass $\Delta \lambda$. The use of filters not only reduces background noise but also improves the signal quality and the efficiency of the PMTs. Filtering background skylight becomes even more significant in II observations, as, currently, these are carried out during full moon nights when the primary function of the IACTs (of observing Cherenkov Showers) is rendered infeasible. It is important to note that the light from the stellar source is focused on a PMT attached with each of the telescopes during II observations.

The significance of the signal can be expressed in terms of the signal-to-noise ratio (SNR), which depends on many factors. However, most importantly, it does not depend on the optical bandwidth $\Delta \nu$ of the radiation for a two-telescope correlation. The explanation for the independence of the SNR from $\Delta \nu$ is provided in several works (e.g., subsection 4.1 of K. N. Rai et al. 2021). The Signal-to-Noise is given by

$$SNR = A \cdot \alpha \cdot q \cdot n \cdot F^{-1} \cdot \sigma \cdot \sqrt{\frac{T \Delta f}{2}} \cdot |V_{12}|^2 \quad (4)$$

Here, A is the total mirror area, α is the quantum efficiency of the PMTs, q is the throughput of the remaining optics, and n is the differential photon flux from the source. The excess noise factor of the PMTs is represented by F , T denotes the observation time, and σ is the normalized spectral distribution of the light (including filters) (e.g., V. A. Acciari et al. 2020).

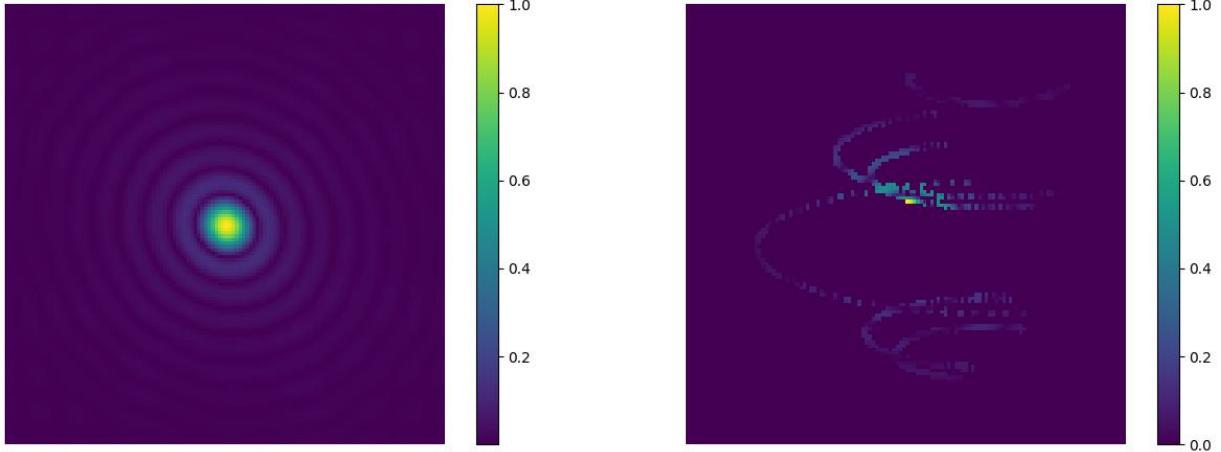


Figure 4. The left panel shows the absolute value of the two-dimensional Fast Fourier Transform of the source depicted in Fig. 3. It represents the intensity interferometric (u, v) plane image of the source that would be obtained by an infinite number of baselines (or an infinite number of telescopes observing the source). The right panel shows the absolute value of the same source measured along the tracks shown in Fig. 2. This figure reflects the sparse nature of the signal received by a realistic finite number of telescopes and baselines sampled from the full (u, v) plane signal space, as shown in the figure on the left panel. Both figures are plotted on a linear scale and normalized to the maximum pixel value in each respective figure.

3.3. Baseline considerations

The measurement of the size of stellar objects via squared visibility depends on the distance between the telescopes, known as the baseline b .

$$|V_{12}(b)|^2 = \frac{c(b)}{c(0)} \quad (5)$$

For achieving a good SNR with a given telescope configuration, covering as much as possible of the interferometric plane is always desirable. If the source is at the zenith, the coordinates in the Fourier plane (u, v) are given by:

$$(u, v) = \frac{1}{\lambda} (b_E, b_N) \quad (6)$$

where b_E and b_N are, respectively, the baselines expressed in east and north coordinates. However, the sources can be anywhere on the sky, and the telescopes are stationary and may also have different relative altitudes b_A depending on the available terrain. In order to gather maximum possible information on the source during the observation session and to cover as much of the observational plane as possible during such sessions, Earth's rotation must be taken into account using rotated baselines. For a given stellar source with declination δ and hour-angle h , as observed by telescopes located at latitude l , equation (7) provides the rotated baselines for a given pair of telescopes (see e.g., eqs. 8–10 from S. Baumgartner et al. 2020) with the R -matrices

representing the respective rotation operations.

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = R_x(\delta) \cdot R_y(h) \cdot R_x(-l) \begin{pmatrix} b_E \\ b_N \\ b_A \end{pmatrix} \quad (7)$$

Fig. 2 shows the track of six baselines generated from the telescopes (Fig. 1) due to the Earth's rotation. Since every pair of telescopes traces an ellipse in the Fourier plane, the total number of ellipses scales as

$$\mathcal{N} = \frac{1}{2} N_T \cdot (N_T - 1) \quad (8)$$

where N_T is the number of telescopes considered. As the number of baselines increases non-linearly, Intensity Interferometry (II) benefits greatly from a large number of telescopes. The CTAO can offer many more baselines — D. Dravins et al. (2013) considered the telescope configurations then being planned and showed how it would provide a dense coverage of the interference plane.

3.4. A Fictitious Fast Rotating Star: Our Test Case

In our work presented here, we attempt image reconstruction of a fast-rotating star using its simulated Intensity Interferometric observation in a cGAN architecture. Fast-rotating stars are important test cases for understanding various astrophysical processes, including stellar evolution, internal structure, and dynamical behaviour over time. Fast rotation causes stars to adopt an oblate shape, flattening at the poles and bulging at the equator due to the stronger centrifugal force (e.g.,

H. Von Zeipel 1924; A. Maeder 1999). Fig. 3 shows an image qualitatively representing a fictitious fast-rotating star, with brightness (and, therefore, the effective surface temperature) distributed across its surface. The brightness (effective temperature) is highest at the poles and lowest at the equator, a phenomenon known as gravity darkening (L. B. Lucy 1967). First direct interferometric detection of stellar photospheric oblateness (of Altair) was pioneered by van Belle et al G. T. van Belle et al. (2001) using the Palomar Testbed Interferometer (PTI) and the Navy Prototype Optical Interferometer (NPOI). Gravity darkening due to fast rotation was first observed through interferometric and spectroscopic data from the CHARA Array for the fast-rotating star Regulus (H. A. McAlister et al. 2005). As pointed out earlier, these two pieces of work, all using Michelson Interferometry, make a subset of several others (G. T. van Belle et al. 2001; A. Domiciano de Souza et al. 2003; H. A. McAlister et al. 2005; A. Domiciano de Souza et al. 2005; J. D. Monnier et al. 2007; E. Pedretti et al. 2009; M. Zhao et al. 2009; A. O. Martinez et al. 2021). The first observation of photospheric oblateness (of γ Cassiopeiae or γ -Cas) using Intensity Interferometry (II) has been recently carried out at VERITAS observatory and is reported by Archer et al. A. Archer et al. (2025). Reportage of such observations of other γ -Cas like targets and other class of FRS by Cherenkov Telescope arrays, such as the MAGIC array, are expected by 2026. In addition, observation and measurement of gravity darkening using II is the natural next step and is yet to be reported. In this context, our work of reconstructing the image of FRSs from their II-simulated observations using cGAN is an attempt at solving this inverse problem along with mitigation of loss of phase information in II.

II counts the photons arriving at the telescopes from the stellar object. The correlation of these photon arrivals at the telescopes yields the squared visibility Eq. (5), as explained in subsection 3.1. The left panel of fig. 4 shows the signal from the source shown in Fig. 3 using II, displayed on linear scales. Point to note here is that this figure represents the signal from the source that would be recorded by an infinite number of baselines provided by an infinite number of telescopes on the interferometric plane. In practice, only a small part of this information is available (as seen in right panel of Fig. 4), because one has a finite number of baselines corresponding to the finite number N_T of telescopes at our disposal and a limited observation schedule. We have simulated the II observation of the fictitious star by a hypothetical observation facility having an array of four IACTs with their relative positions approximating

those at VERITAS (correlated with baselines as seen in Fig. 1) over one night. Using this modest amount of signal from one night's observation, we have trained a cGAN to construct the image of the source.

4. GENERATIVE ADVERSARIAL NETWORKS

Generative Adversarial Networks (GANs) were introduced by I. Goodfellow et al. (2014). A GAN model involves two competing deep neural network models, referred to as the Generator and the Discriminator. These two networks engage in a zero-sum “Minimax” game, as in the Game Theory. Given a real data set $\{x_i\}$ (for example, a set of real images) drawn from some unknown distribution $P_{\text{data}}(x)$ generated by some unknown or ill-understood process, the objective here is to generate a new set (of images) whose probability distribution should match $P_{\text{data}}(x)$ as closely as possible. During the training of the two models, the Generator samples a latent variable z from a known prior distribution $P_z(z)$ (e.g., the Normal Distribution) and produces a synthetic sample $G(z)$ to start with and, subsequently, based on updates received from the Discriminator as its training progresses. The Discriminator, being a probabilistic binary classifier, receives either a real data sample x or a generated sample $G(z)$, and outputs a probability that the input is real. The Discriminator aims to maximise its classification accuracy, while the Generator aims to fool it by trying to minimize it. The training of these two networks proceeds alternately leading to the optimization of the adversarial loss function $L(D, G)$ given by

$$\begin{aligned} L(D, G) &= \min_G \max_D V(D, G) \\ &= \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] \\ &\quad + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))] . \end{aligned} \quad (9)$$

The two neural networks $G(z) \equiv G(z; \theta_G)$ and $D(x) \equiv D(x; \theta_D)$ are parameterized by θ_i with $i = G$ or D respectively. During its training, the Generator generates the differentiable function $x_{\text{gen}} \equiv G(z)$ that maps z to the data space x . Through such maps the Generator generates its own distribution $p_G(x_{\text{gen}})$ and through the training episodes, specifically by iteratively updating its parameters θ_G , aligns this distribution with the distribution of real data $p_{\text{data}}(x)$. The training data set provided to the Discriminator is constructed by mixing real data points x and generated data points x_{gen} in equal ratio. The Discriminator generates the function $D(x)$ that represents the probability that x originates from real data. Eq.(9) implies that training of the GAN model moves towards maximization of the expectation of $D(x)$. Through this process, both the parameters θ_G

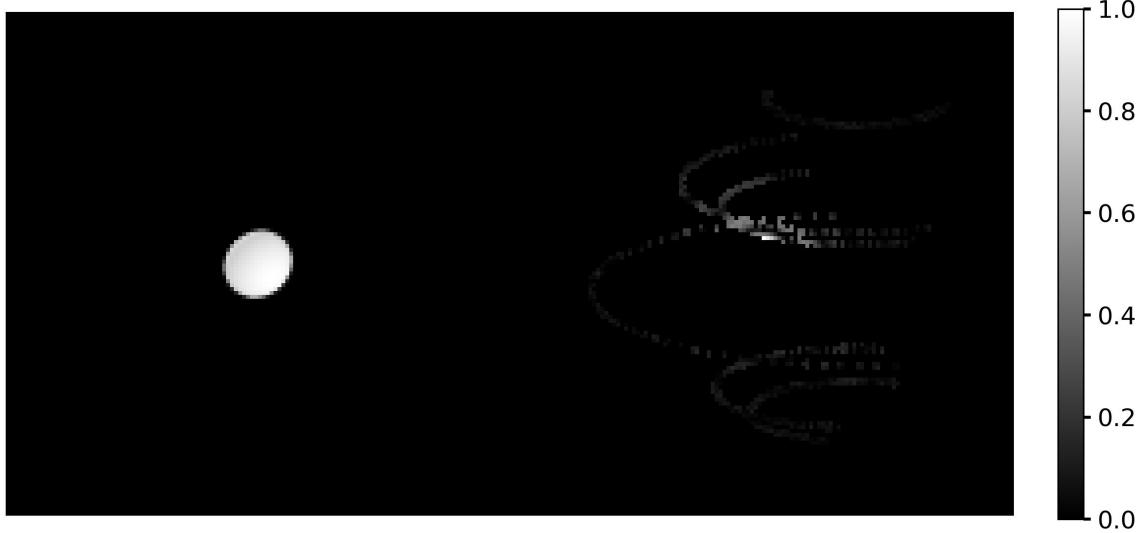


Figure 5. An illustrative example of the input used for training the cGAN model. The picture on the left shows the source image, which serves as the ground truth or the real data (x), as mentioned in the Flow Diagram (Fig. 6) discussed later during the training). The picture on the right represents the simulated II observation pattern of the source (on the left) using a hypothetical observation facility having as array of four IACTs (see Fig. 1) and the tracks of the six telescope baselines (see Fig. 2) generated by these four IACTs due to Earth’s rotation during the observation session. This pattern referred to as y , in the Flow Diagram (Fig. 6) discussed later) forms the “condition” during the training to which the GAN model has to conform. Salt-and-pepper noise is added to this pattern for enhancing the robustness of the cGAN model. Together, these images form a training pair, where the GAN learns to reconstruct a predicted image (modeling of observed signal) similar to ground truth (left) from the noisy baseline signal (right). The grayscale in both images is normalized to the brightest pixel.

and θ_D are optimized such that $p_G(x_{gen})$ gets maximally aligned with $p_{data}(x)$.

For a given fixed Generator $G(z)$, the problem can be reformulated as:

$$\max_D V(D, G) = \mathbb{E}_{x \sim p_{data}} [\log D_G^*(x)] + \mathbb{E}_{x \sim p_G} [\log (1 - D_G^*(x))] \quad (10)$$

where D_G^* denotes the optimum of the discriminator. As seen in equation (11), the global optimum of equation (10) is achieved if and only if $p_G = p_{data}$. Furthermore, if both G and D are allowed to reach their respective optima, – the so called Nash point of the Minimax game – p_G converges to p_{data} .

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_G(x_{gen})} \quad (11)$$

At this point, the Discriminator finds its job no better than random guessing. A more comprehensive discussion of the problem, including proofs, is provided in I. Goodfellow et al. (2014).

Subsequently, the GAN framework was extended to a conditional model (M. Mirza et al. 2014). This new model, known as “conditional GAN” or cGAN injects a conditioning variable y into both networks: the Generator now generates $G(z | y)$, and the Discriminator

evaluates $D(x | y)$. The adversarial objective becomes

$$V(D, G) = \mathbb{E}_{x, y \sim p_{data}(x)} [\log D(x|y)] + \mathbb{E}_{z, y \sim p_z(z)} [\log (1 - D(G(z|y)|y))] \quad (12)$$

The conditional variable y in a cGAN can be various additional information including images, labels or text contextual to “ground truth” real data x . Among various types of cGANs, Pix2Pix GAN with its image-to-image translation design is suitable for the task at hand. In our work, we specifically use this conditional variable by choosing y to represent the ground-based intensity-interferometry (II) observation patterns: the Generator is trained to produce stellar surface images that not only look realistic but also conform to the measured II correlations, while the Discriminator judges realism *and* consistency with the II data.

P. Isola et al. (2017) further observed that combining the cGAN from Eq. (12) with the traditional $L_1(G)$ loss (also known as the Mean Absolute Error) improves the results, as the Generator is encouraged to produce outputs closer to the target image in a pixel-wise sense. We adopt this approach in the training of our cGAN model by including $L_1(G)$ defined in eq.(13)

$$L_1(G) = \mathbb{E}_{x \sim p_{data}, z \sim p_z(z)} [\|x - G(z | y)\|_1]. \quad (13)$$

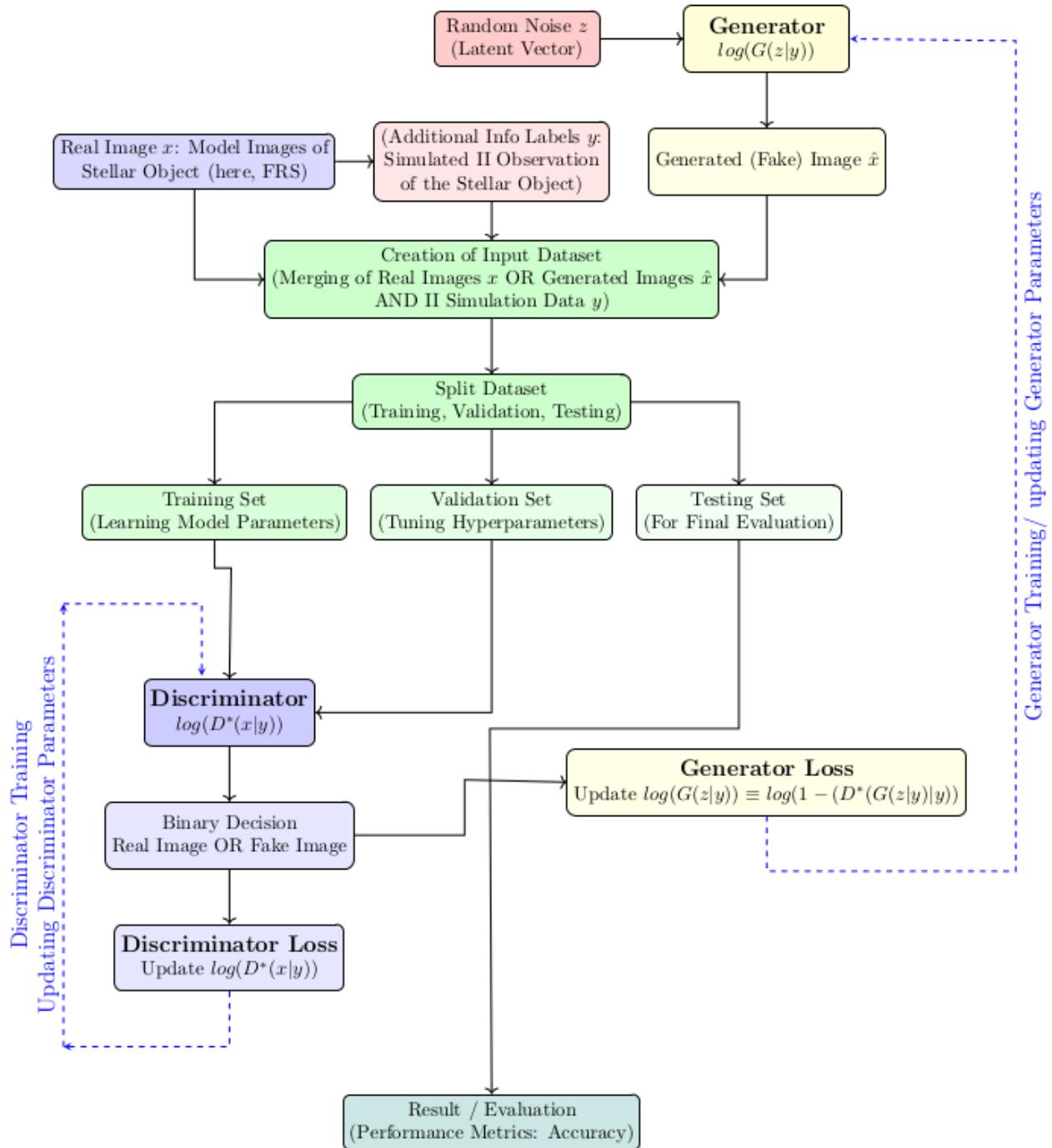


Figure 6. A schematic representation of the features of a cGAN model used in this work and the iterative process of its training, validation and testing. The process constitutes four broad stages: (1) choice of an appropriate GAN architecture including both the Discriminator and the Generator (not shown in this figure) (2) preparation of the Training, Validation and Testing datasets and (3) Training and Validation of the Model (4) Testing and Evaluation of the Model. The stages (2), (3) and (4) are depicted in this figure. The datasets are prepared in three broad steps: (i) generating the “ground truth” images of fictitious fast-rotators x , the sparse II images y used as the “condition” images in the Model and the generated images z sampled from a Normal Distribution (ii)merging these images into individual files with $(x|y)$ and/or $(z|y)$ (as seen in an illustrated sample in Fig.(5)) and generating the full dataset in this process, and finally (iii) partitioning the full dataset into Training Set, Validation Set and the Testing Set. After the iterative training of the Model and its validation process is complete (“Nash point” of the Minimax Game is reached), the Model is tested using the Testing Set and evaluated.

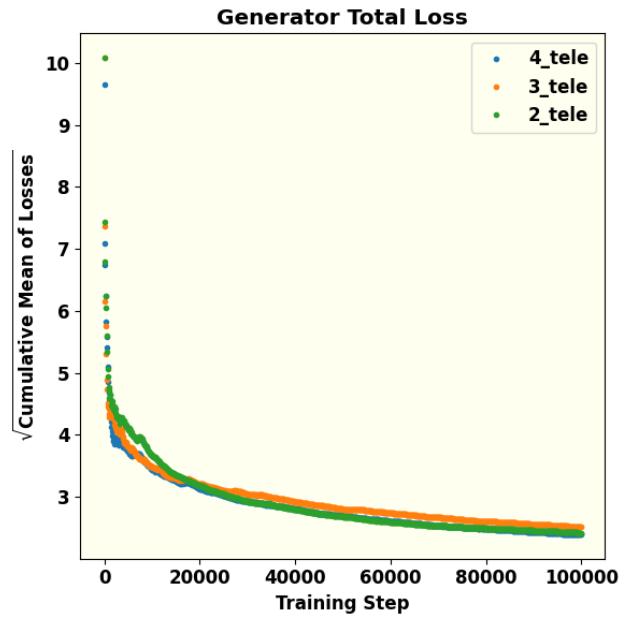
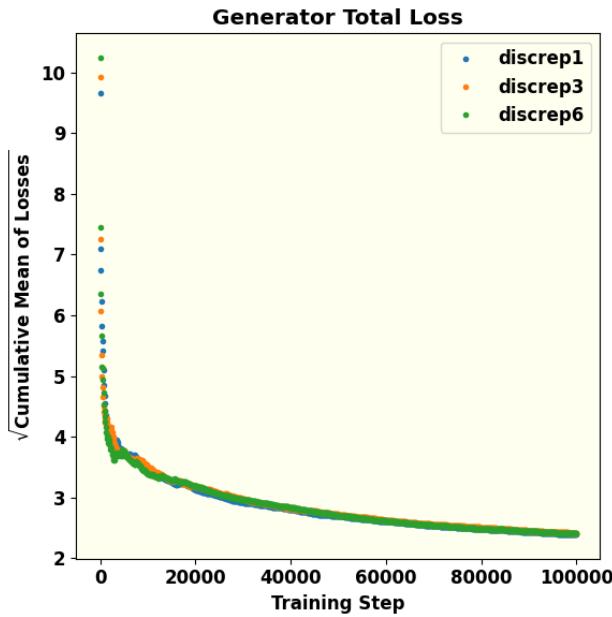
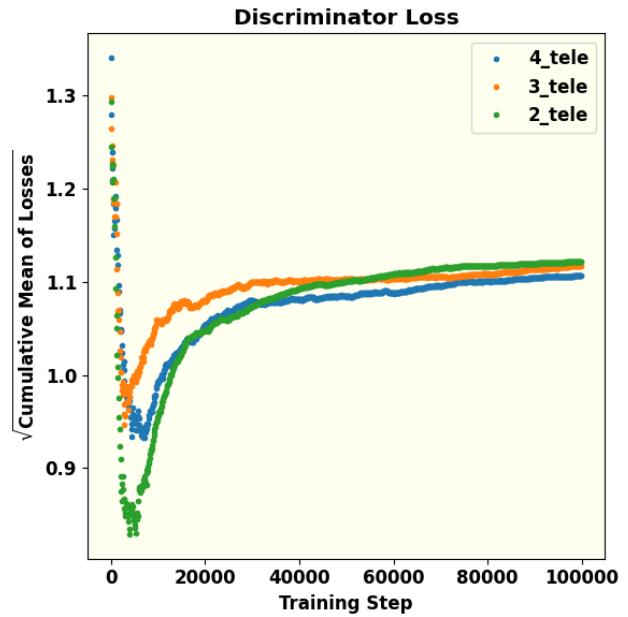
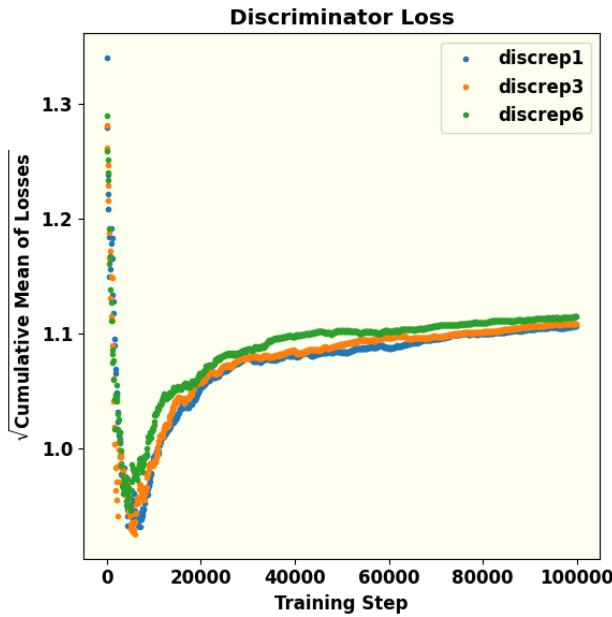


Figure 7. These figures show the effect of the ratio of episodes of Discriminator training per every episode of Generator training. This hyperparameter is termed Discriminator repetition or (discrep) in the figures. The square root of the cumulative mean of the losses are plotted against the training steps for 3 values of this ratio. Understandably, this ratio has a higher impact on the Discriminator loss than the Generator loss. Equal number of episodes of training produces minimum cumulative loss of the Discriminator. The dip in the Discriminator loss during the initial phases of its training can be interpreted as its early success in detecting “fake” (or generated) images because of a poorly trained Generator. With gradual training of the Generator, the success rate of the Discriminator decreases and eventually approaches saturation with equal probability of being successful in telling “fake” from real. The continual decrease and eventual saturation of the Generator loss is a result of its training to generate better images with increasing number of steps.
 DOI: [10.5281/zenodo.1759867](https://doi.org/10.5281/zenodo.1759867)

Figure 8. The square root of cumulative mean of Discriminator and Generator loss for different numbers of telescopes. The number of telescopes is also another hyperparameter that has significant impact on the model performance. If there are only two telescopes, both Discriminator and Generator are not trained smoothly. The result of four telescopes is a lot better because the cumulative mean of loss functions is smaller compare to other parameters. For the same reasons as explained in Fig.(7), we observe initial dip and eventual saturation of the Discriminator loss and continual decrease and eventual saturation of the Generator loss with training steps.

500 The total adversarial loss function, along with the L_1
 501 loss modulated by a hyperparameter λ then becomes

$$502 \quad L_{tot} = \arg \min_G \max_D V(D, G) + \lambda \cdot L_1(G). \quad (14)$$

503
 504 This type of network has demonstrated remarkable ro-
 505 bustness across a variety of applications. For example, it
 506 can generate colored images from grayscale inputs based
 507 on architectural labels, transform images from day to
 508 night, and even predict maps from satellite data. A more
 509 extensive list of applications is provided in P. Isola et al.
 510 (2017).

511 4.1. Generator

512 As discussed above, in a GAN the Generator, a deep
 513 neural network in itself, is responsible for producing
 514 synthetic data, in this case, images that resemble those
 515 of a fast-rotating star. In this work, the Generator is
 516 implemented as a U-Net convolutional network (O. Ron-
 517 neberger et al. 2015). The U-Net consists of a symmet-
 518 ric encoder-decoder structure forming a characteristic
 519 “U” shape along with skip connections: the left (con-
 520 tracting or down-sampling) path, the right (expanding
 521 or up-sampling path) and the connecting (bottle-neck)
 522 path. The left down-sampling path repeatedly applies
 523 3×3 convolutions (padded to preserve spatial dimen-
 524 sions) followed by LeakyReLU activations and strided
 525 convolution (with stride of 2), by progressively halv-
 526 ing the spatial resolution while doubling the channel
 527 depth (typically $64 \rightarrow 128 \rightarrow 256 \rightarrow 512 \rightarrow 1024$). At
 528 the bottleneck, high-level features are processed with-
 529 out further spatial reduction. The right up-sampling
 530 path mirrors this process using 2×2 transposed convolu-
 531 tions (stride 2) for up-sampling, halving the channel
 532 count at each level and using ReLU as the activation
 533 function for all its layers except the output layer. Ad-
 534 ditionally, a dropout layer is introduced at the begin-
 535 ning of the upsampling phase to mitigate overfitting of
 536 the Generator model (P. Isola et al. 2017). Critically,
 537 long skip connections concatenate feature maps from
 538 each encoder (down-sampling) level to the correspond-
 539 ing decoder (up-sampling level), directly injecting high-
 540 resolution details into the reconstruction process. This
 541 enables the network to “remember where everything is”
 542 while the deep bottleneck still provides the large recep-
 543 tive field needed to think globally about image structure
 544 and semantics. Besides the strided convolutions, mod-
 545 ern variants of U-Nets used in state-of-the-art GANs
 546 incorporate residual blocks within resolution levels, and
 547 frequently add spectral normalization and self-attention
 548 at the bottleneck for improved training stability and
 549 long-range dependency modelling. These architectural

550 choices allow our U-Net generator to simultaneously re-
 551 cover sharp, high-frequency details (stellar limb edges,
 552 limb-darkening profiles, gravity darkening, and rapid-
 553 rotation-induced oblateness) and enforce global coher-
 554 ence (overall disk morphology and physical consistency
 555 with the observed interferometric visibilities)—making
 556 it ideally suited for high-fidelity sky-image reconstruc-
 557 tion of fast-rotating stars from sparse ground-based in-
 558 tensity interferometry observations.

559 During the training, the total Generator Loss func-
 560 tion L_G including the L_1 loss defined in eq.(13) that is
 561 minimized is given by

$$562 \quad L_G = -\mathbb{E}_{z \sim p_z(z)} [\log D(G(z | y) | y)] \\ + \lambda \mathbb{E}_{x \sim p_{\text{data}}, z \sim p_z(z)} [\|x - G(z | y)\|_1]. \quad (15)$$

563 Here,

- 564 • x denotes a real data sample (e.g. the “ground-
 565 truth” image; here, the synthetically generated
 566 fast rotator image) corresponding to condition y ,
 567 the simulated II observation data.
- 568 • z is a random latent vector drawn from the prior
 569 $p_z(z)$.
- 570 • $G(z | y)$ is the generator output (the reconstructed
 571 / synthesized image) given z and condition y .
- 572 • $D(\cdot | y)$ is the discriminator’s estimate (probabil-
 573 ity) that its input is “real,” given the same condi-
 574 tion y .
- 575 • λ is a hyperparameter controlling the trade-off be-
 576 tween adversarial realism and pixel-wise fidelity
 577 (typical values depend on the problem, e.g. in
 578 pix2pix, $\lambda = 100$).

579 4.2. Discriminator

580 The Discriminator is tasked with classifying the im-
 581 ages produced by the Generator as either real or fake.
 582 It takes a real image from the dataset (often referred
 583 to as the target image for the Generator) and provides
 584 feedback to guide the Generator toward producing more
 585 accurate images. In this work, the PatchGAN model (P.
 586 Isola et al. 2017) is employed as the Discriminator. Un-
 587 like a traditional global classifier, PatchGAN evaluates
 588 individual patches of the image, outputting a grid of
 589 predictions rather than a single scalar value. Each el-
 590 ement in the grid corresponds to the “realness” of one
 591 patch of the image under examination of the Discrimi-
 592 nator at a time. The final loss of the Discriminator is
 593 the average over all the patch responses. Evaluating the
 594 “realness” of the input image in terms of its constituent
 595 patches facilitates capture of texture/ style and other

high frequency components in the image. As compared to a global discriminator, it also reduces the number of parameters in the network thereby helping reduce computation cost. It also works on images with arbitrary sizes.

Prior to the down-sampling of data using PatchGAN, each input image is preprocessed with application of Zero Padding followed by batch normalization. The purpose of Zero Padding is to prevent the loss of spatial information and to facilitate the extraction of deeper features from the down-sampled output. Batch normalization is required to stabilize the learning (loss minimization) process. PatchGAN then reduces the spatial dimensions of the images to extract localized features, ensuring the model focuses on smaller regions. In this down-sampling stage, a leaky version of the Rectified Linear Unit (LeakyReLU) is applied in the convolutional layers, similar to the approach used in the Generator. The probability $D(\cdot|y)$ that the patch of the input image represented by \cdot is “real” is evaluated through this process. The loss function D of the full input image is obtained by averaging over all the patch responses.

The Discriminator Loss function L_D that is optimized during the training process is given by

$$L_D = -\mathbb{E}_{x \sim p_{\text{data}}(x|y)} [\log D(x|y)] - \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z|y)|y))] \quad (16)$$

where the arguments of the D and G functions are as noted in the text following eq.(15). This loss is composed of two parts: one that assesses how accurately it identifies fake images (by comparing predictions to a target value of 0) and the other that measures how accurately the Discriminator identifies real images (by comparing predictions to a target value of 1).

The training procedure of these two components of the cGAN model can be outlined as follows:

- The discriminator D is updated by minimizing L_D , keeping G fixed (so that D becomes better at distinguishing real vs generated images).
- The generator G is updated by minimizing L_G , keeping D fixed, thus pushing G to generate images that are (a) judged “real” by D , and (b) close (in pixel-wise sense) to the ground-truth x .

5. NETWORK PARAMETERS

An appropriate cGAN architecture along with a set of hyperparameters was optimized by tuning. The objective of the model, as already mentioned, was to learn to faithfully reproduce a set of sky-images of fast rotators subject to the condition that those images are consistent with their simulated II observation data. We discuss

below the architecture and the hyperparameters of the cGAN model used for this task. Given the adversarial nature of GANs, where the Generator and Discriminator engage in a minimax game, careful tuning of key parameters is critical to ensure that both networks are well-balanced for effective training.

5.1. Data Preparation

First, we generate synthetic images of rapidly rotating stars by modelling them as oblate spheroids with varying radii and oblateness parameters between 0.5 and 1. To incorporate the effect of gravity darkening, we also consider different viewing angles and assume a linear dependence on the declination angle of each point on the stellar surface. The traced ellipses result from integrating over the source’s hour angle.

Next, Salt and Pepper noise is introduced into these images; usually this is done at the rate of 0.5% of the number of pixels in the image. Then, the images are resized and their mean is subtracted. A two-dimensional Fast Fourier Transform, along with a Fourier shift, is applied, yielding a complex number for each pixel. Since II does not measure phase, the absolute value is calculated.

Next, sampling of the interferometric plane is introduced via pixel-wise multiplication of the absolute-valued Fourier-transformed image (left panel of Fig. 4) and the baseline tracks of (Fig. 2) generated by the four IACTs due to Earth rotation during the II observation. The result is a map in the Fourier plane featuring several ellipses, which is also referred to as the sparse sampling map (right panel of fig. 4). This map represents the sparse sampling of the signal space corresponding to the source (Fig. 3) observed with four telescopes (Fig. 1). In essence, the right panel of fig. 4 represents the result of simulated II observation of the fictitious fast-rotator illustrated in fig. 3.

Finally, we normalize the pixel values and convert them to 8-bit integers, producing an image that encodes the sparsely sampled, phase-free visibility measured by II. The image of the corresponding simulated (fictitious) star, which serves as the “ground truth” is processed identically to avoid any bias. These two images are merged side-by-side into a single image (as shown in Fig. 5). The dataset thus created is split into three parts in the proportions of 80% for training, 10% for validation, and 10% as the test set. This partition of the full dataset is indicated in the Fig. 6. In the following, we refer to these three parts as the Training Set, the Validation Set and the Testing Set respectively.

5.2. GAN Architecture

The GAN architecture used in this work is a Pix2Pix cGAN, which uses an image-to-image translation strat-

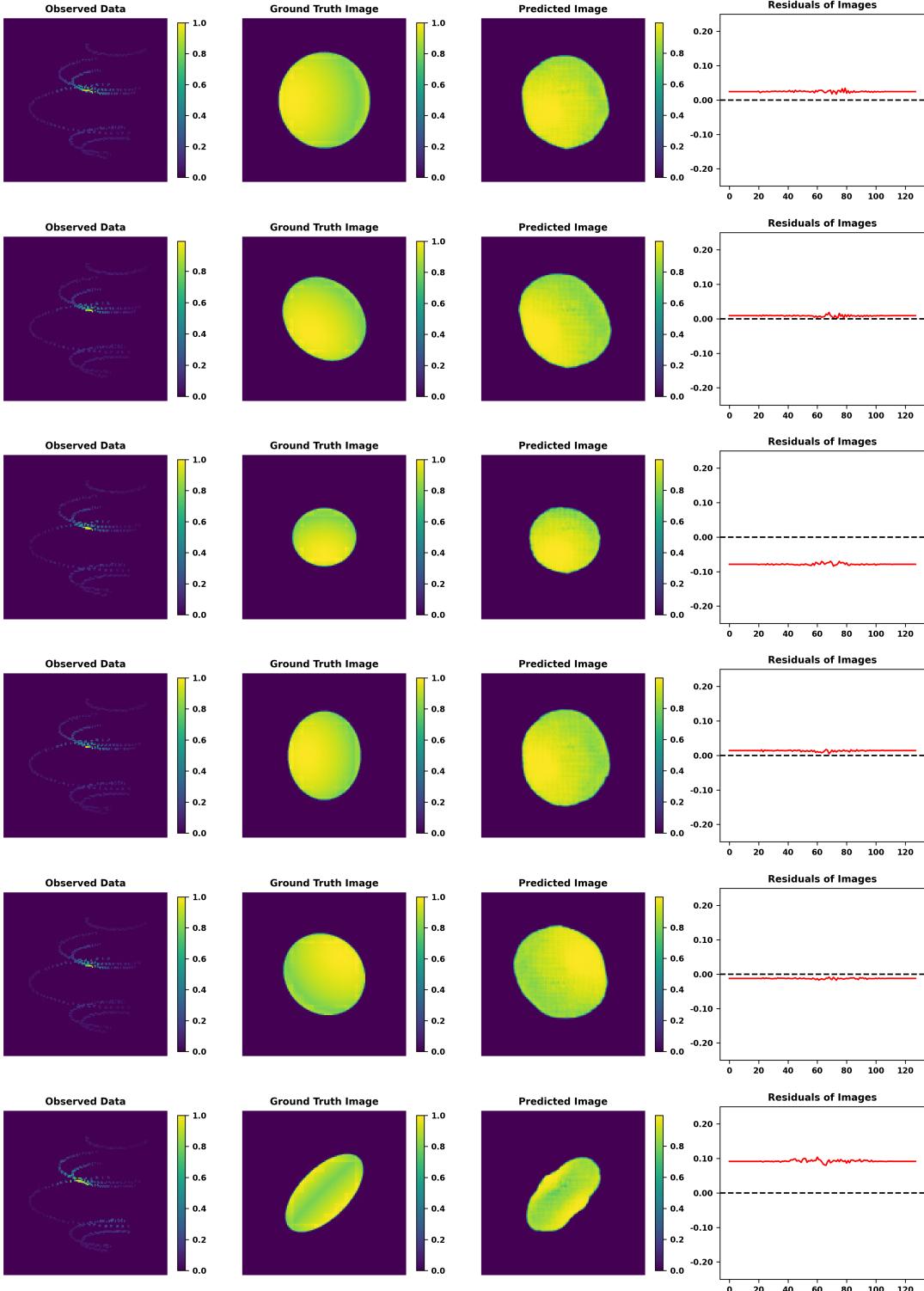


Figure 9. Example results of image reconstruction using the cGAN model along with the II observations simulated in this work. Each row in this figure represents the results for a hypothetical fast-rotating star. The left panel represents the sparse II pattern obtained by the simulated observation of the star using the four IACTs illustrated in Fig.1. This image acts as the “condition” part of the data input to the cGAN model. The second panel from left displays the real image, or ground truth, which the Discriminator uses to distinguish from the images generated by the Generator. The data generated for training, validation and testing of the cGAN model is a merge of this image and its II pattern presented in the left panel. The third panel is the reconstructed image, or the predicted image, produced by the trained GAN model . The fourth and the last panel is the difference between the ground truth and the predicted image in the (u, v) plane.

egy with both the ground truth and the condition being images. Originally introduced by Isola et al. P. Isola et al. (2017), this architecture is specifically suitable for image processing and reconstruction objectives. For instance, the TensorFlow tutorials⁶ demonstrate its application to a dataset of architectural facades. This architecture has been adapted for the phase retrieval problem at hand here. The network is implemented using the TensorFlow library (M. Abadi et al. 2016), calculations are performed with `scipy` (P. Virtanen et al. 2020), and plots are generated with `matplotlib` (J. D. Hunter 2007).

5.3. Hyperparameter Tuning

The cGAN model architecture used in this work employs several hyper-parameters, which are explained briefly below (for a more in-depth discussion, see K. P. Murphy 2022).

The learning rate (lr) of the optimizer determines how much the model updates its parameters with each iteration. A learning rate that is too small may lead to underfitting, while one that is too large can render the model unstable. Therefore, selecting an appropriate learning rate is crucial (K. P. Murphy 2022). A canonical choice in Pix2Pix and other GAN models is $lr = 2 \times 10^{-4}$. In our case too, we found this choice to be appropriate.

The kernel size refers to the dimensions of the convolutional kernel used in the network, determining how many pixels are combined to produce a new pixel. A larger kernel size can capture features spanning several pixels, but it may also incorporate unrelated features. Small kernel sizes are preferable in cases where target images finer details or high spatial frequency features. Since the “ground truth” target images in our case have longer scale gravity darkening features, we have opted for the more canonical choice of kernel size being 5.

The amount of noise is controlled by two parameters, “alpha” and “beta”, which indicate the percentage of pixels altered to either white or black, hence the term Salt and Pepper noise. Here, “alpha” is applied to the real image, while “beta” is applied to the generated image. Different ratios (“alpha/beta”) can lead to varying model performance; however, our results indicate that distinct noise rates do not significantly affect the loss functions.

The batch size defines the number of images processed simultaneously by the network. Smaller batch sizes have been observed to improve generalization (S. J. Prince 2023). However, because a larger batch size significantly increases training time, a batch size of 1 is used. Besides,

in Pix2Pix cGAN implementations found in literature this choice is found to be often preferred.

Buffer size of a Pix2Pix GAN refers to a small memory of image pool of previously generated images. They are occasionally fed to the Discriminator in place of the freshly generated ones. This strategy of mixing old and new fakes mitigates the risk of mode collapse wherein the Discriminator tends to map all or large number of generated images to only one or a few real “ground truth” image(s). We have chosen a fairly large buffer size (=1400) to protect the model against mode collapse.

In the training of GANs, one often-followed strategy to potentially boost performance is to give the Discriminator an advantage by increasing its number of training epochs before returning to the Generator’s training. This hyperparameter is referred to as the Discriminator repetition (as seen in Table 1). While this can lower the Discriminator loss, as shown in Fig. 7, it also increases training time. In training our model, we did not notice any significant advantage derived from this strategy. Since the generated images did not noticeably improve with additional Discriminator training, we adopted the strategy of training both the networks with equal preference (Discriminator repetition = 1).

One domain specific hyperparameter of the cGAN model presented here is the Number of Telescopes N_T . The degree of sparse sampling of the intensity interferometric (II) image plane can be varied to provide the model with access to more number of active (non-zero) pixels. Point to note here, is that the coverage of the Fourier interferometric plane (number of active pixels in the II image) scales with available number of baselines, and the latter scales quadratically with the number of telescopes. Fig. 8 shows the loss functions for different numbers of telescopes. There is a significant disparity in performance of the model: at the “Nash point”, both the loss functions are minimized for the case of four telescopes. Overall, the degree of sparse sampling appears to have the most pronounced effect of all the hyperparameters.

The hyperparameter Output Channels refers to the number of channels in the generator output (e.g., 1 for grayscale, 3 for RGB). It is worthwhile to recall that the cGAN model constructed in this work is trained on “ground truth” target images and the simulated II data, both in grey scales as seen in Fig. 5. It is natural that the output of this model will be in grey scales only. Therefore the value of this hyperparameter is set to 1. The choice of hyperparameter λ has been commented upon earlier.

An optimized set of hyperparameters is selected through an iterative process of training and validation.

⁶ <https://www.tensorflow.org/tutorials/generative/pix2pix>

For a tentatively chosen set of the hyperparameters, the model is trained using the Training Set until the both the Discriminator and the Generator loss functions are minimized. This model, thus trained along with its model parameters, is then passed through validation using the Validation Set. This cycle of training and validation is iterated till an optimal set of hyperparameters is arrived at. During each epoch of training of the model, both the Discriminator and the Generator networks are trained for 100,000 steps. Plots of the "Discriminator Loss" function and the "Generator Total Loss" function presented in Fig. 7 and Fig. 8 represent the results of this training for two of the hyperparameters, namely, the Discriminator repetition ("discrep", in short) and the Number of telescopes, respectively. Obviously, the most compute-intensive part of this process is that of the training of the Model. The results of training and validation presented in this work were carried out on a CPU using two nodes, each with 48 threads and the entire process of training and validation required approximately 20 hours on the machine employed for this work. This iterative process of training and validation of the Model is represented schematically in the Fig. 6. The chosen optimal set of hyperparameters is presented in the table Table-1. This optimized and trained Model

Table 1. Selected hyperparameters used for training the model.

Hyperparameter	Value
Learning rate	2e-4
Kernel size	5
Alpha/Beta	1
Batch size	1
Buffer Size	1400
Discriminator repetition	1
Number of Telescopes	4
Output Channels	1
Lambda	100

is then subjected to testing and evaluation using the Testing Set. The results of this testing and evaluation is presented in the following section.

The Pix2Pix cGAN architecture along with the choice of the values of the hyperparameters mentioned in the Table -1 and used in the implementation of this architecture constitutes the cGAN model (hereinafter referred to as "the GAN Model" or simply "the Model").

6. IMAGE RECONSTRUCTION: RESULTS AND EVALUATION

In this section, we examine the performance of the GAN Model whose architecture and choice of hyperpa-

rameters have been discussed above. We subject the trained Model to the task of phase retrieval and image reconstruction on the Testing Dataset.

6.1. Visual Evaluation of Images Predicted by the Model

Fig. 9 demonstrates the success of the trained Model in reconstructing the images of a sample of the fictitious fast-rotators drawn from the Testing Set.

The four panels from left to right in each row of Fig. 9 represent the following:

- The left panel represents the sparse II pattern obtained by the simulated observation using the four IACTs illustrated in Fig. 1. As explained earlier, this image acts as part of the input, namely its "condition" part. This image acts as the condition that the sky-images of the star generated by the Generator must conform to.
- The second panel from left displays the real image, or ground truth target image, which the Discriminator loss function uses to distinguish from the images generated by the Generator.
- The third panel from left presents the reconstructed (or predicted) image corresponding to the ground truth (second panel) and generated by the trained Model. The similarity of these two images indicates the success of the Model in its stated objective of image reconstruction.
- The right panel shows the residuals between the ground truth target image and the predicted image in the interferometric plane. The small values are indicative of the success of training the Model.

The predicted images and the residuals presented in the third and the fourth columns (from left) of Fig. 9 show encouraging results, conveying visual information about the source's size, shape, and brightness distribution across its surface fairly accurately. This has been achieved on the basis of the input provided by II observation using only six baselines. Arguably, further improvements can be achieved by increasing the number of telescopes to enhance the coverage of the (u, v) plane. A closer examination of this proposition might be able to contribute to the design and instrumentation aspects in the existing and upcoming CTAO.

6.2. Evaluation of the Model using Moments

The reconstructed images are visually compelling, demonstrating the Model's effectiveness in using II to reconstruct images. However, visual assessment alone is

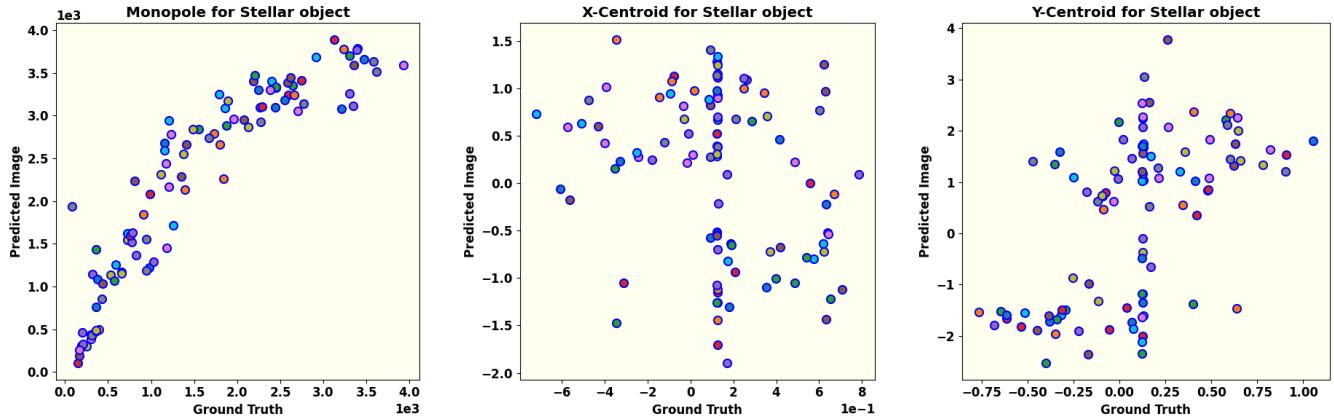


Figure 10. This set of figures shows comparisons of the monopole, and the coordinates of the centroids (x_c, y_c) of the pixel brightness distribution of the ground truth images and the corresponding images predicted (generated) by the trained model. The monopole moment of the images represents the overall pixel brightness of the images. The values, as expected of a trained Model, not only lie in same range, but lie close to the diagonal and show an approximate equality. The centroids of both the ground truth images and the predicted images lie within small ranges of pixel values as should be expected of an ideally trained model. Please see the relevant text for detailed comments.

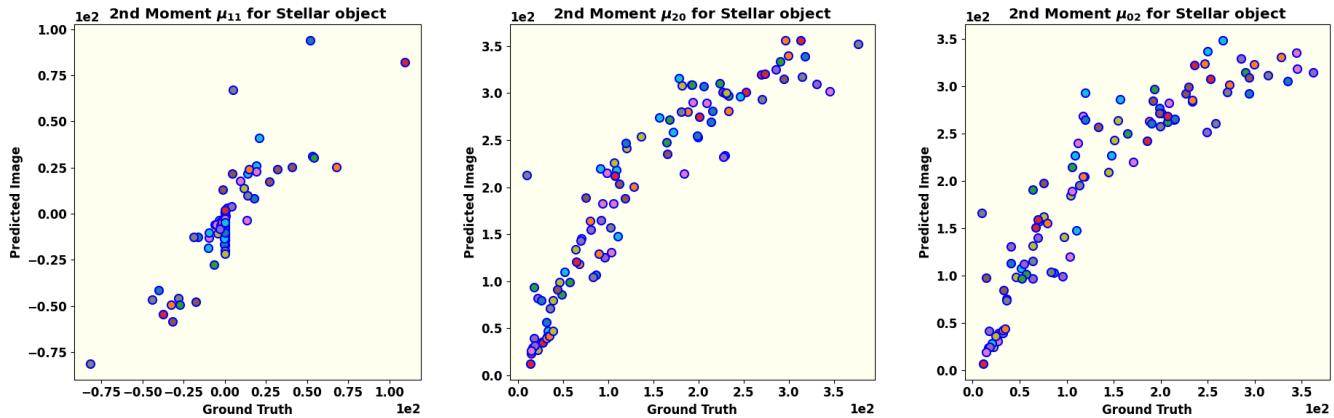


Figure 11. The second order central moment of the brightness distribution of the images is analogous to the moment of inertia of a mass distribution. The panels of this figure show the comparison of the second order moments (μ_{11}, μ_{20} , and μ_{02}) of the predicted images vs. the ground truth images from left to right respectively. Approximate equality of these moments is evident. The small scatter in the moments of the predicted images is indicative of a balanced training of the Model.

insufficient; statistical evaluation of the generated images in comparison with the ground truth images is necessary to validate the results. We present here the results of our calculation and comparison of the moments of distribution pixel brightness in the target “Ground Truth” images and the “Predicted” images.

Image moments capture key properties of the reconstructed objects, such as shape, size, and intensity distribution, by quantifying features like position, orientation, and brightness distribution. By comparing the moments of the Model-generated images with those of the ground truth, we can objectively assess the consistency and accuracy of the reconstruction. This approach provides a reliable framework for evaluating reconstruction

quality, as image moments can reveal subtle differences in geometric and intensity properties that might not be apparent through visual inspection alone.

The raw moment M_{ij} of an image is defined as (M.-K. Hu 1962)

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (17)$$

where $I(x, y)$ represents the intensity at pixel (x, y) . The zeroth order raw moment, or monopole, represents the total intensity of an image. It is computed by summing all pixel values across the image, yielding an overall intensity measure. In this context, analyzing the monopole provides the total pixel brightness of the images of the fictitious stars. According to Eq. (17), the

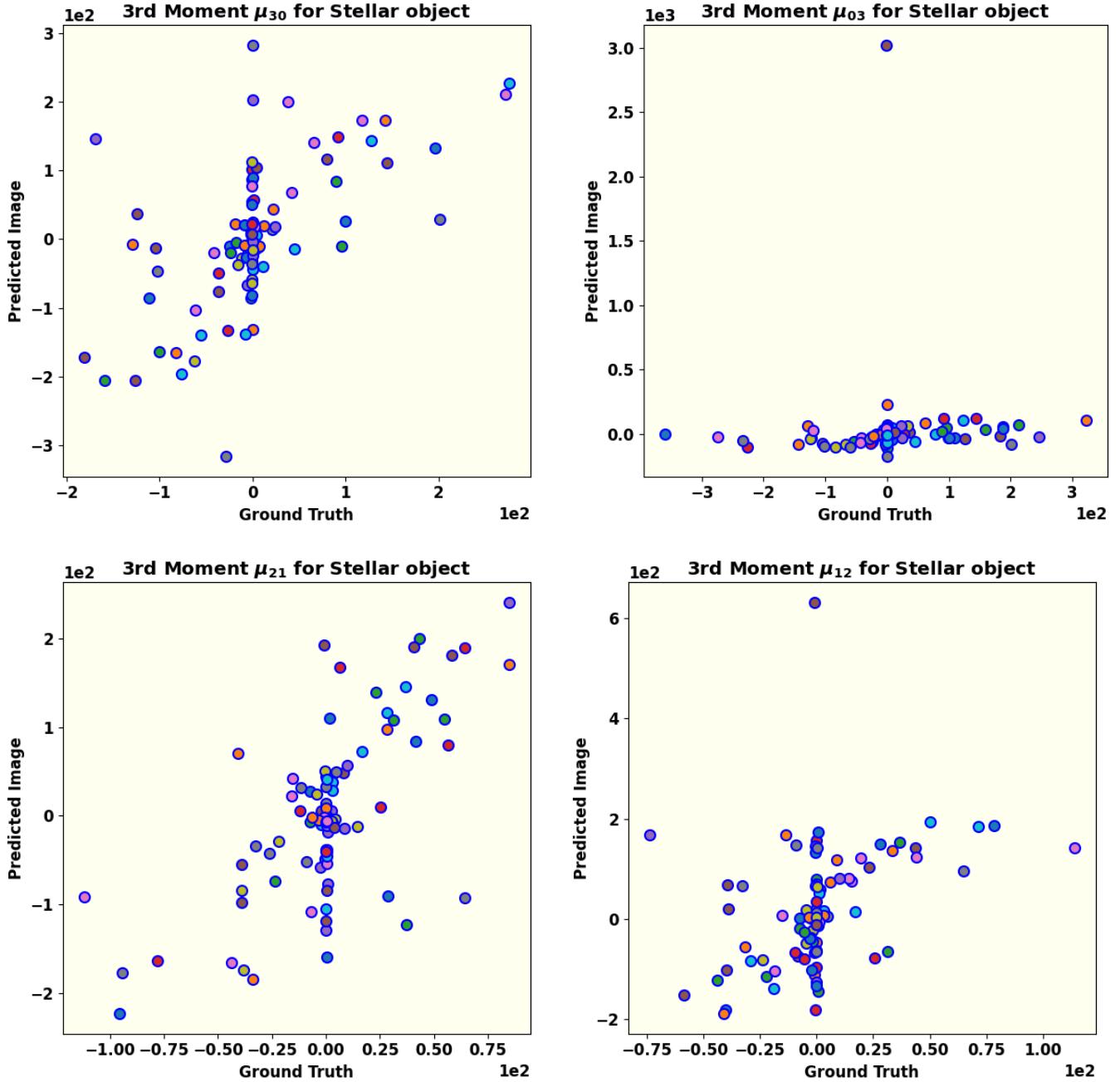


Figure 12. Shown here are all the third-order central moments (μ_{30} , μ_{03} , μ_{21} , and μ_{12}) for ground truth and predicted images generated by the trained Model. They represent the skewness of the brightness distributions. The low skewness in the brightness distribution of the ground truth images is evident. One outlier among the predicted images dominates their skewness distribution along the y -axis as seen in the values of μ_{03} and μ_{12} . The trained Model, however, seems to have learned the approximately uniform distribution of skewness along the x -axis as is evident, with a few outliers though, in the values of μ_{30} and μ_{21} .

908 monopole of an image is calculated as:

$$M_{00} = \sum_x \sum_y I(x, y). \quad (18)$$

910
911 An ideally trained Model should predict the overall
912 pixel brightness of the generated images to be in the same range of values, if not equal to those of the ground
913 truth images. The left panel of Fig. 10 displays the
914 monopole values for 100 reconstructed images. The plot
915 reveals a broad agreement in the overall intensity of the
916 ground truth targets and the predicted (generated) im-
917 ages. However, a slightly higher brightness of recon-
918 structed images in the intermediate range of brightness
919 is evident. This could, perhaps, be mitigated by tuning
920 the Batch Size hyperparameter.

921 While the monopole effectively represents the total
922 brightness, it does not provide information about the position,
923 shape, size, or detailed brightness distribution of the fast-rotating stars. For these aspects, higher-order
924 moments are necessary.

925 The first order raw moments normalized by the re-
926 spective monopole moments of the images given by

$$\begin{aligned} x_c &= \frac{M_{10}}{M_{00}} = \frac{\sum_{x,y} x \cdot I(x, y)}{\sum_{x,y} I(x, y)} \\ y_c &= \frac{M_{01}}{M_{00}} = \frac{\sum_{x,y} y \cdot I(x, y)}{\sum_{x,y} I(x, y)} \end{aligned} \quad (19)$$

931 represent the centroids (x_c, y_c) of the pixel brightness
932 distribution of the images. An ideally trained GAN
933 model should produce the centroids of the target and
934 predicted images within a small range of pixels. A larger
935 range would imply a low fidelity in learning the bright-
936 ness distribution of ground truth images. The middle
937 and right panels of Fig. 10 compare the centroids (x_c, y_c)
938 of 100 predicted images with their corresponding ground
939 truths respectively. We notice that the coordinates of
940 the centroids of ground truth images are spread across a
941 range of 2 pixels whereas the those of the predicted im-
942 ages are spread across a span of 4 pixels along x -axis and
943 7 pixels along y -axis. This clustering of centroids within
944 a small scale range across all results indicates that the
945 reconstructed images represent the spatial distribution
946 of brightness of predicted and ground truth images fairly
947 accurately.

948 Furthermore, these calculated centroids are instru-
949 mental in analyzing the shape, size, and brightness dis-
950 tribution of the stars using higher-order image moments.
951 To this end, the central moment of an image is calcu-
952 lated according to:

$$\mu_{pq} = \frac{1}{M_{00}} \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y). \quad (20)$$

953 The sum of p and q defines the order of the central
954 moment.

955 The second order central moment of the brightness
956 distribution of the images is analogous to the moment of
957 inertia of a mass distribution. Fig. 11 presents the com-
958 parison of second-order central moments $(\mu_{11}, \mu_{20}, \mu_{02})$,
959 which are used to study the structure of a fast-rotating
960 star along the line of sight (as explained in the upcom-
961 ing subsection). All three plots demonstrate an approx-
962 imate equality among these moments, similar to the
963 monopole, thereby confirming the success of applying
964 the Model to reconstruct images with II. The small scat-
965 ter in the moments of the predicted images indicates
966 robust learning of the Model without over-fitting.

967 The third-order central moments $(\mu_{30}, \mu_{03}, \mu_{21}, \mu_{12})$ of
968 the images quantify the degree and direction of asymme-
969 try of the brightness distribution around their respective
970 centroids. Fig. 12 presents a comparison of all third-
971 order moments for both the ground truth and the recon-
972 structed image. All the four panels show a low skewness
973 in the brightness distribution of the ground truth im-
974 ages. The skewness of the predicted images along the
975 y -axis is dominated by one outlier as seen in the values of
976 μ_{03} and μ_{12} . However, the approximately uniform dis-
977 tribution of skewness of the predicted images along the
978 x -axis, with a few outliers though, as seen in the values
979 of μ_{30} and μ_{21} suggests that this feature in the ground
980 truth images has been picked up fairly well during the
981 training of the Model.

983 6.3. The reconstructed Parameters for object

984 The centroids (x_c, y_c) indicate only the center of the
985 star and its spatial location in the image. In contrast,
986 the second-order central moments determine the ori-
987 entation, semi-major axis, and eccentricity relative to the
988 source's center (M. R. Teague 1980). These moment-
989 based parameters fully describe the two-dimensional el-
990 lipse that fits the image data.

991 The orientation of a fast-rotating star along the line of
992 sight is defined in terms of second-order central moments
993 as

$$\theta = \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right). \quad (21)$$

995 The semi-major and semi-minor axes of the stellar ob-
996 ject are computed using the second-order central mo-
997 ments and are denoted as a and b , respectively.

$$\begin{aligned} a &= 2\sqrt{mp + \delta} \\ b &= 2\sqrt{mp - \delta} \end{aligned} \quad (22)$$

999 where,

$$mp = \frac{\mu_{20} + \mu_{02}}{2} \quad (23)$$

1001 and

$$\delta = \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{2}. \quad (24)$$

1003 Using the calculated axis values, the eccentricity of the
1004 fast-rotating star is determined as:

$$e = \sqrt{1 - a/b}. \quad (25)$$

1006 Eqs. 21–25 describe the elliptical nature of the stellar
1007 object (in this case, a fast-rotating star) and provide in-
1008 formation on its shape and size, depending on the com-
1009 puted values. In contrast, the brightness distribution
1010 is characterized by skewness, which is quantified using
1011 third and higher-order moments.

7. CONCLUSION

1013 Intensity Interferometry (II) is re-emerging as a
1014 promising technique to overcome the challenges of very
1015 long baseline interferometry in the optical wavelength
1016 range. However, compared to radio-interferometry, opti-
1017 cal interferometry faces a major hurdle: photon corre-
1018 lation captures only the magnitude of the interferometric
1019 signal, resulting in a loss of phase information.

1020 This work addresses the challenge of phase retrieval in
1021 II using a machine-learning technique, specifically a con-
1022 ditional Generative Adversarial Network (cGAN). Our
1023 study demonstrates that the size, shape and brightness
1024 distribution of fast rotating stars can be recovered by
1025 a Pix2Pix cGAN model trained on the combined input
1026 of the sky-image of known sources along with their re-
1027 spective II data. In this training the sky-image acts
1028 as the real “ground truth” and the II data acts as the
1029 “condition”. The Discriminator of our Model is trained
1030 to efficiently distinguish between the real images and
1031 fake (generated) images based on the “ground truth”
1032 images and the respective II data as the condition. The
1033 Generator is trained to generate progressively realistic
1034 images which are also consistent with the condition of
1035 the II data. The evaluation of the trained Model is then
1036 carried out by comparison of image moments of ground

1037 truth images and generated images. Specifically, the
1038 monopole, second, and third-order moments are com-
1039 pared. The results support the effectiveness of cGAN in
1040 achieving the stated objective.

1041 While the results of this study highlight the significant
1042 potential of machine learning, and in particular the ap-
1043 plicability of cGAN, for image reconstruction in II, sev-
1044 eral aspects require further refinement. First, an impor-
1045 tant factor in the reconstruction process is the extent of
1046 Fourier plane coverage, which depends on the number of
1047 available telescopes and the total observing time. The
1048 reasonable success of this piece of work suggests that a
1049 network of higher number of telescopes providing higher
1050 number of baselines and greater coverage of the (u, v)
1051 plane signal, would play a critical role in projects of im-
1052 age reconstruction of more complicated stellar systems
1053 can be undertaken. Future work might explore different
1054 observatory layouts to assess their impact on image re-
1055 construction quality. Second, detector efficiencies, which
1056 impact the signal-to-noise ratio (SNR) of actual observa-
1057 tional data, have not yet been incorporated; addressing
1058 these factors will be crucial for more accurate SNR es-
1059 timation. Third, exploring and comparing alternative
1060 methods for image generation could reveal approaches
1061 that outperform cGAN in reconstructing stellar images
1062 with II. Fourth, experimenting with different loss func-
1063 tions could provide additional insights into the recon-
1064 struction quality. Although further testing is needed to
1065 refine the GAN and enhance its robustness and reli-
1066 ability, our findings suggest that machine learning is a
1067 promising approach for phase reconstruction in II.

ACKNOWLEDGEMENTS

1068
1069 **Acknowledgment:** One of the authors (SS) gratefully
1070 acknowledges the computing facilities and the local hos-
1071 pitality extended to him by the Inter-University Centre
1072 for Astronomy and Astrophysics (IUCAA), Pune, In-
1073 dia under its Visiting Associate Programme during the
1074 preparation and finalization of this manuscript.

REFERENCES

- 1075 Abadi et al., M. 2016, arXiv preprint arXiv:1603.04467
- 1076 Abe et al., S. 2024, MNRAS, 529, 4387
- 1077 Acciari et al., V. A. 2020, MNRAS, 491, 1540
- 1078 Acharyya et al., A. 2024, The Astrophysical Journal, 966,
- 1079 28
- 1080 Aleksić et al., J. 2016, Astroparticle Physics, 72, 76
- 1081 Archer, A., Aufdenberg, J. P., Bangale, P., et al. 2025,
1082 arXiv e-prints, arXiv:2506.15027,
1083 doi: [10.48550/arXiv.2506.15027](https://doi.org/10.48550/arXiv.2506.15027)
- 1084 Baumgartner et al., S. 2020, MNRAS, 498, 4577
- 1085 Coccomini et al., D. 2021, arXiv preprint arXiv:2122.11578
- 1086 Domiciano de Souza, A., Kervella, P., Jankov, S., et al.
1087 2003, Astronomy and Astrophysics, 407, L47,
1088 doi: [10.1051/0004-6361:20030868](https://doi.org/10.1051/0004-6361:20030868)

- 1089 Domiciano de Souza, A., Kervella, P., Jankov, S., et al.
 1090 2005, *Astronomy and Astrophysics*, 442, 567,
 1091 doi: [10.1051/0004-6361:20042476](https://doi.org/10.1051/0004-6361:20042476)
- 1092 Dravins et al., D. 2013, *Astroparticle Physics*, 43, 331,
 1093 doi: [10.1016/j.astropartphys.2012.04.017](https://doi.org/10.1016/j.astropartphys.2012.04.017)
- 1094 Fienup, J. 1982, *Applied Optics*, 21, 2758
- 1095 Gamo, H. 1963, *Journal of Applied Physics*, 34, 875
- 1096 Gerchberg, R. W. 1972, *Optik*, 35, 237
- 1097 Glauber, R. J. 1963, *Physical Review*, 130, 2529
- 1098 Goldberger et al., M. L. 1963, *Physical Review*, 132, 2764
- 1099 Goodfellow et al., I. 2014, *Advances in neural information
 1100 processing systems*, 27
- 1101 Hanbury Brown, R., & Twiss, R. Q. 1956, *Nature*, 177, 27
- 1102 Hanbury Brown et al., R. 1957, *Proceedings of the Royal
 1103 Society of London. Series A. Mathematical and Physical
 1104 Sciences*, 242, 300
- 1105 Hanbury Brown et al., R. 1958, *Proceedings of the Royal
 1106 Society of London. Series A. Mathematical and Physical
 1107 Sciences*, 243, 291
- 1108 Hanbury Brown et al., R. 1974, *Monthly Notices of the
 1109 Royal Astronomical Society*, 167, 121
- 1110 Haubois et al., X. 2009, *Astronomy & Astrophysics*, 508,
 1111 923, doi: [10.1051/0004-6361/200912927](https://doi.org/10.1051/0004-6361/200912927)
- 1112 Holmes et al., R. 2010, in *Adaptive Coded Aperture
 1113 Imaging, Non-Imaging, and Unconventional Imaging
 1114 Sensor Systems II*, Vol. 7818, SPIE, 175–185
- 1115 Hu, M.-K. 1962, *IRE transactions on information theory*, 8,
 1116 179
- 1117 Hunter, J. D. 2007, *Computing in Science & Engineering*, 9,
 1118 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- 1119 Isola et al., P. 2017, in *Proceedings of the IEEE conference
 1120 on computer vision and pattern recognition*, 1125–1134
- 1121 Kirisits et al., C. 2024, arXiv preprint arXiv:2406.14143v2
- 1122 Le Bohec, S.; Holder, J. 2006, *The Astrophysical Journal*,
 1123 649, 399
- 1124 Li et al., X. 2014, in *Imaging Spectroscopy, Telescopes and
 1125 Large Optics*, Vol. 9298, SPIE, 92981G1–G7
- 1126 Liu, L.-C., Wu, C., Li, W., et al. 2024, arXiv preprint,
 1127 arXiv:2404.15685, doi: [10.48550/arXiv.2404.15685](https://doi.org/10.48550/arXiv.2404.15685)
- 1128 Liu et al., L.-C. 2025, *Physical Review Letters*, 134, 180201,
 1129 doi: [10.1103/PhysRevLett.134.180201](https://doi.org/10.1103/PhysRevLett.134.180201)
- 1130 Lucy, L. B. 1967, *Zeitschrift für Astrophysik*, Vol. 65, p. 89,
 1131 65, 89
- 1132 Maeder, A. 1999, *Astronomy & Astrophysics*, 347, 185
- 1133 Martinez, A. O., Baron, F., Monnier, J. D., Roettenbacher,
 1134 R. M., & Parks, J. R. 2021, *The Astrophysical Journal*,
 1135 916, 60, doi: [10.3847/1538-4357/ac06a5](https://doi.org/10.3847/1538-4357/ac06a5)
- 1136 McAlister et al., H. A. 2005, *The Astrophysical Journal*,
 1137 628, 439
- 1138 Mirza et al., M. 2014, arXiv preprint arXiv:1411.1784
- 1139 Monnier, J. D., Zhao, M., Pedretti, E., et al. 2007, *Science*,
 1140 317, 342, doi: [10.1126/science.1143205](https://doi.org/10.1126/science.1143205)
 DOI: [10.5281/zenodo.17598807](https://doi.org/10.5281/zenodo.17598807)
- 1141 Murphy, K. P. 2022, *Probabilistic machine learning: an
 1142 introduction* (MIT press)
- 1143 Mustafa et al., M. 2019, *Computational Astrophysics and
 1144 Cosmology*, 6, 1
- 1145 Norris et al., R. P. 2021, *The Astrophysical Journal*, 919,
 1146 124, doi: [10.3847/1538-4357/ac0c7e](https://doi.org/10.3847/1538-4357/ac0c7e)
- 1147 Nuñez et al., P. D. 2010, in *Optical and Infrared
 1148 Interferometry II*, Vol. 7734, SPIE, 458–467
- 1149 Nuñez et al., P. D. 2012a, *Monthly Notices of the Royal
 1150 Astronomical Society*, 419, 172
- 1151 Nuñez et al., P. D. 2012b, *Monthly Notices of the Royal
 1152 Astronomical Society*, 424, 1006
- 1153 Pedretti, E., Monnier, J. D., ten Brummelaar, T., &
 1154 Thureau, N. D. 2009, *New Astronomy Reviews*, 53, 353,
 1155 doi: [10.1016/j.newar.2010.07.005](https://doi.org/10.1016/j.newar.2010.07.005)
- 1156 Prince, S. J. 2023, *Understanding deep learning* (MIT
 1157 press)
- 1158 Rai, K. N., Basak, S., Sarangi, S., & Saha, P. 2025,
 1159 *Resonance*, 30, 45, doi: [10.1007/s12045-025-1729-x](https://doi.org/10.1007/s12045-025-1729-x)
- 1160 Rai et al., K. N. 2021, *Monthly Notices of the Royal
 1161 Astronomical Society*, 507, 2813,
 1162 doi: [10.1093/mnras/stab2391](https://doi.org/10.1093/mnras/stab2391)
- 1163 Rai et al., K. N. 2022, *Monthly Notices of the Royal
 1164 Astronomical Society*, 516, 2864,
 1165 doi: [10.1093/mnras/stac2433](https://doi.org/10.1093/mnras/stac2433)
- 1166 Ronneberger et al., O. 2015, in *Medical image computing
 1167 and computer-assisted intervention—MICCAI 2015: 18th
 1168 international conference, Munich, Germany, October 5–9,
 1169 2015, proceedings, part III* 18, Springer, 234–241
- 1170 Sato et al., T. 1978, *Applied optics*, 17, 2047
- 1171 Sato et al., T. 1979, *Applied Optics*, 18, 485
- 1172 Sato et al., T. 1981, *Applied Optics*, 20, 2055
- 1173 Schawinski et al., K. 2017, *MNRAS*, 467, L110
- 1174 Teague, M. R. 1980, *Josa*, 70, 920
- 1175 Teague, M. R. 1983, *Journal of Optical Society of
 1176 America*, 73, 1434
- 1177 van Belle, G. T., Ciardi, D. R., Thompson, R. R., Akeson,
 1178 R. L., & Lada, E. A. 2001, *The Astrophysical Journal*,
 1179 559, 1155, doi: [10.1086/322340](https://doi.org/10.1086/322340)
- 1180 Virtanen et al., P. 2020, *Nature methods*, 17, 261
- 1181 Vogel et al., N. 2025, *MNRAS*, 537, 2334
- 1182 Von Zeipel, H. 1924, *Monthly Notices of the Royal
 1183 Astronomical Society*, Vol. 84, p. 665–683, 84, 665
- 1184 Zhang et al., J. 2020, *Opt. Lett.*, 45, 3649
- 1185 Zhao, M., Monnier, J. D., Pedretti, E., et al. 2009, *The
 1186 Astrophysical Journal*, 701, 209,
 1187 doi: [10.1088/0004-637X/701/1/209](https://doi.org/10.1088/0004-637X/701/1/209)