Deep Horizon

A machine learning network that recovers accreting black hole parameters

van der Gucht, J.; Davelaar, J.; Hendriks, L.; Porth, O.; Olivares, H.; Mizuno, Y.; Fromm, C.M.; Falcke, H.

Citation for published version (APA):

Astronomy
&
Astrophysics

Deep Horizon: A machine learning network that recovers accreting black hole parameters

Jeffrey van der Gucht[1], Jordy Davelaar[1,2], Luc Hendriks[1], Oliver Porth[4,3], Hector Olivares[3], Yosuke Mizuno[3], Christian M. Fromm[3,5], and Heino Falcke[1]

[1] Department of Astrophysics/IMAPP, Radboud University, PO Box 9010, 6500 GL Nijmegen, The Netherlands
e-mail: j.davelaar@astro.ru.nl
[2] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[3] Institut für Theoretische Physik, Max-von-Laue-Straße 1, 60438 Frankfurt am Main, Germany
[4] Anton Pannekoek Instituut, Universiteit van Amsterdam, PO Box 94249, 1090 GE Amsterdam, The Netherlands
[5] Max-Planck Institute for Radio Astronomy, Auf dem Huegel 69, 53115 Bonn, Germany

ABSTRACT

Context. The Event Horizon Telescope recently observed the first shadow of a black hole. Images like this can potentially be used to test or constrain theories of gravity and deepen the understanding in plasma physics at event horizon scales, which requires accurate parameter estimations.
Aims. In this work, we present `Deep Horizon`, two convolutional deep neural networks that recover the physical parameters from images of black hole shadows. We investigate the effects of a limited telescope resolution and observations at higher frequencies.
Methods. We trained two convolutional deep neural networks on a large image library of simulated mock data. The first network is a Bayesian deep neural regression network and is used to recover the viewing angle i, and position angle, mass accretion rate $\dot{M}$, electron heating prescription R_{high} and the black hole mass M_{BH}. The second network is a classification network that recovers the black hole spin a.
Results. We find that with the current resolution of the Event Horizon Telescope, it is only possible to accurately recover a limited number of parameters of a static image, namely the mass and mass accretion rate. Since potential future space-based observing missions will operate at frequencies above 230 GHz, we also investigated the applicability of our network at a frequency of 690 GHz. The expected resolution of space-based missions is higher than the current resolution of the Event Horizon Telescope, and we show that `Deep Horizon` can accurately recover the parameters of simulated observations with a comparable resolution to such missions.

Key words. accretion, accretion disks – black hole physics – radiative transfer – methods: data analysis

1. Introduction

In April 2019, the Event Horizon Telescope (EHT) collaboration released the first image of the shadow of a black hole (Event Horizon Telescope Collaboration 2019a,b,c,d,e,f). This image is direct evidence of the existence of black holes, a fundamental prediction of the general theory of relativity (GR; Schwarzschild 1916; Kerr 1963).

In GR, astrophysical black holes are characterized by their mass, M_{BH}, and their spin, $a = Jc/GM_{BH}^2$. In this equation, J is the angular momentum of the black hole, G is the gravitational constant, and c is the speed of light. The size of the black hole is set by its event horizon, $R_h = (1 + \sqrt{1 - a^2})R_g$, where $R_g \equiv GM_{BH}/c^2$ is the gravitational radius. The event horizon defines a surface from within nothing can escape. The event horizon is gravitationally lensed, resulting in an effective angular size of $\theta = 2\sqrt{27}R_g/D$ for $a = 0$, where D is the distance to the black hole. This lensed image is known as the shadow (Falcke et al. 2000). Although the scale of the observed shadow is on the order θ, the exact size depends on both the emission model and GR effects, such as spacetime rotation. Therefore, models of the accretion flow around black holes are needed to interpret the results of the EHT (Event Horizon Telescope Collaboration 2019e).

The EHT array consists of eight telescopes positioned all around the globe, resulting in an effective resolution of ~20 microarcseconds (μas) when operating at 1.3 mm (Event Horizon Telescope Collaboration 2019b). With this effective resolution, the EHT resolved the shadow of the black hole M 87*, that is, a supermassive black hole (SMBH) in the nucleus of Messier 87. The distance to this SMBH is 16.8 ± 0.8 Mpc (Bird et al. 2010; Cantiello et al. 2018) and the mass is $6.5 \pm 0.7 \times 10^9\ M_\odot$ (Event Horizon Telescope Collaboration 2019f). This mass estimate is computed from the observed angular size on the sky of $42 \pm 3\ \mu$as (Event Horizon Telescope Collaboration 2019a).

The size of the Earth limits the resolution of the EHT. Furthermore, the EHT only sparsely samples the Fourier domain of the image (*uv*-plane) (Event Horizon Telescope Collaboration 2019b), owing to the limited amount of suitable millimeter Very Long Baseline Interferometry (VLBI) telescope sites. Increasing the amount of coverage in the *uv*-plane increases the quality of the image. Both of these limitations are mitigated by switching to space-based VLBI (SVLBI). Furthermore, SVLBI would remove atmospheric corruption and allow for longer baselines and higher frequencies. Therefore, SVLBI allows for higher resolutions and improved image quality, compared to ground-based VLBI. There are several studies of future SVLBI missions that observe the shadow of a black

hole (Palumbo et al. 2018; Fish et al. 2020; Roelofs et al. 2019). Roelofs et al. (2019) report simulations of future SVLBI measurements of the black hole shadow of Sagittarius A*, the SMBH in our galaxy, up to a frequency of 690 GHz. Their setup has baselines up to 60 Gλ, resulting in a resolution of 4 μas after several months of observations. Density fluctuations in the interstellar medium electrons cause phase fluctuations in the incoming plane wave, resulting in scattering of the radio wave (Narayan & Goodman 1989; Goodman & Narayan 1989; Johnson & Gwinn 2015). At a frequency of 690 GHz, there is less interstellar scattering (Bower et al. 2006; Roelofs et al. 2019), and the measured emission originates from closer to the event horizon as compared to the EHT observations.

The image of a black hole shadow can be used to test and constrain theories of gravity (Johannsen & Psaltis 2010; Psaltis et al. 2015; Goddi et al. 2017; Event Horizon Telescope Collaboration 2019f; Mizuno et al. 2018), but this requires accurate parameter estimations. Previous studies of M 87* often use general relativistic magnetohydrodynamical (GRMHD) simulations to model the accretion flow (Dexter et al. 2012; Mościbrodzka et al. 2016, 2017; Ryan et al. 2018; Davelaar et al. 2018, 2019; Chael et al. 2019a). These studies fit their models to the observed spectra, resulting in constraints on the model parameters. In Event Horizon Telescope Collaboration (2019e,f), the model parameters are constrained by fitting GRMHD models to the image of M 87*. The appearance of the black hole shadow in the image is determined by the parameters and can, therefore, be recovered directly from the image. In the case of EHT (Event Horizon Telescope Collaboration 2019e,f), GRMHD models were compared with the data. The models are either MAD or SANE, and include five spin values. The image library was then constructed based on these GRMHD models by performing general relativistic ray-tracing (GRRT) simulation for six values of the temperature ratio of electrons to protons inside the accretion disk, parametrized by $R_{\rm high}$ (Mościbrodzka et al. 2016, 2017), one mass, and two inclinations. The images were scaled in post-processing to test for other masses as well (e.g., not generated directly with ray-tracing codes). The current scoring of the GRMHD/GRRT images is conducted either via the single snapshot method (SSM; in Event Horizon Telescope Collaboration 2019e) or via the average image scoring (AIS; in Event Horizon Telescope Collaboration 2019f). Both approaches are performed in Fourier space, and a χ^2 between the data and the model is computed using the visibility amplitude and closure phase. During the fitting, the model images are rescaled (in flux density), rotated, and stretched (changing the mass to distance ratio). Currently, two pipelines– THEMIS using Markov chain Monte Carlo (MCMC; Broderick et al. 2020) and GENA, using evolutionary algorithms (Fromm et al. 2019) –are used to perform the fitting of the GRRT images. These pipelines require either MCMC steps, in the case of THEMIS, or generations, in the case of GENA, to provide matching between the data and the images. The results from this comparison, however, show that comparing images to the data results in almost all models fitting the EHT 2017 data (see Table 2, Col. 4 in Event Horizon Telescope Collaboration 2019e); models are mainly rejected based on uncertain measurements of the jet power of M 87 at larger wavelengths (and scales). To both improve the extracting of black hole parameters and to decrease computational needs, we performed a proof of concept to use machine learning for this vital task. In recent years, machine learning algorithms have shown to be efficient and accurate in various fields of astrophysics, including in galaxy classification (Odewahn et al. 1992; Weir et al. 1995; Suchkov et al. 2005; Ball et al. 2006; Vasconcellos et al. 2011; Fadely et al.

2012; Sevilla-Noarbe & Etayo-Sotos 2015; Kim et al. 2015; Kim & Brunner 2017; Lukic & Brüggen 2017), gravitational wave parameter analysis (George & Huerta 2018; Shen et al. 2019; Fan et al. 2019), asteroseismology (Bellinger et al. 2016; Hon et al. 2017; Hendriks & Aerts 2019), and gravitational lensing effects (Hezaveh et al. 2017; Perreault Levasseur et al. 2017; Petrillo et al. 2017; Jacobs et al. 2017).

In this paper, we present Deep Horizon, two Bayesian convolutional deep neural networks that can accurately recover the input parameters of an image of the shadow of an accreting black hole. This network was constructed as a proof of concept to investigate if deep neural networks are capable of obtaining black hole models parameters from horizon scale images. In this proof of concept, we focus on six parameters: the viewing angle with respect to the black hole spin axis, i, mass accretion rate, $\dot{M}$, temperature ratio of electrons to protons inside the accretion disk, mass of the black hole, $M_{\rm BH}$, position angle (PA), and spin of the black hole, a. Our neural network also returns a Bayesian motivated uncertainty on the parameter estimations of the first five parameters mentioned above. We use synthetic images to train and test our neural network, and we restrict ourselves to a single SMBH, M 87*, for which we adopt a distance $D = 16.4$ Mpc; this is slightly smaller (by 2%) than the value used in Event Horizon Telescope Collaboration (2019e), but since it is a general scale factor for the total emission recorded it does not affect the results of this proof of concept. The data are generated at two frequencies, 230 GHz (EHT) and 690 GHz (SVLBI). Furthermore, we investigate the effects of convolving our images with a Gaussian beam as an approximation of a limited telescope resolution (Event Horizon Telescope Collaboration 2019e).

We organize the paper as follows: in Sect. 2 we describe our synthetic data generation and the machine learning methods. In Sect. 3 we show the performance of Deep Horizon on mock observations. In Sect. 4 we discuss our results and future improvements. In Sect. 5 we summarize our results.

2. Methods

Machine learning is a data-driven approach that requires sufficiently large data sets to train the algorithm. In the problem treated in this paper, observational data are limited. Hence, we have to rely on current simulations to generate mock observations of the environment near a black hole.

We generated two data sets, each consisting of 100.000 images. Only the frequency varies between the two sets. The images are computed by post-processing five different GRMHD simulations. The GRMHD data are generated with the Black Hole Accretion Code (BHAC; Porth et al. 2017, 2019; Olivares et al. 2019)[1], and the post-processing is done via the GRRT code RAPTOR (Bronzwaer et al. 2018)[2].

2.1. GRMHD simulations

The two relevant physical parameters of a GRMHD simulation are the spin a and the absolute magnetic flux Φ through the horizon often used in the dimensionless form $\phi = \Phi / \sqrt{\dot{M} R_{\rm g}^2 c}$ (Tchekhovskoy et al. 2011; Porth et al. 2019). In this paper, we only consider standard and normal evolution (SANE; Narayan et al. 2012) models with $\phi \sim 1$, and we use models with a

---

[1] Publicly availble at https://bhac.science
[2] Publicly availble at https://github.com/tbronzwaer/raptor

spin of $a = 0$, ± 0.5 and ± 0.9375. These simulations are part of the simulation library that is used in Event Horizon Telescope Collaboration (2019e) and are initialized with a weakly magnetized Fishbone-Moncrief torus (Fishbone & Moncrief 1976) in orbit around the black hole. The thermal pressure is perturbed with white noise to initialize the magnetorotational instability (MRI). The MRI causes angular momentum to be transported, triggering accretion onto the black hole. The differential rotation of the spacetime and magnetic field lines causes a magnetized jet to launch.

2.2. GRRT simulations

To calculate mock observations, we post-processed the GRMHD data with the GRRT code RAPTOR. This code calculates the flux density map at a given frequency by computing null geodesics, starting from a virtual camera, and simultaneously performing radiative transport calculations. We used emission and absorption coefficients for thermal synchrotron emission. The RAPTOR code used the "fast light" paradigm, where the simulation is frozen with respect to the elapsed photon time, which is equivalent to infinite light speed. We computed images at 230 GHz and 690 GHz. We used a camera with a field of view of (0.1×0.1) milli-arcsec^2 and generated images at (128×128) pixels.

The GRMHD simulations are scale-free. Therefore, we had to convert the GRMHD variables from code units to centimeter-gram-second (cgs) units. This is done by defining the simulation length unit $\mathcal{L} = R_g$, the simulation time unit $\mathcal{T} = R_g/c$, and the simulation mass unit $\mathcal{M}$, where $\mathcal{M}$ sets the density in the accretion flow. The dimensionless accretion rate $\dot{M}_{sim}$ can be converted into the accretion rate in cgs units by $\dot{M} = \dot{M}_{sim}\mathcal{M}/\mathcal{T}$. The variables M_{BH} and $\dot{M}$ are varied in our data generation.

The GRMHD simulation does not evolve the radiatively important electrons. We used a parametrization for the plasma variables, which is based on the assumption that the proton-to-electron coupling depends on plasma magnetization (Mościbrodzka et al. 2016, 2017; Davelaar et al. 2018, 2019; Event Horizon Telescope Collaboration 2019e). This coupling is described by the following formula:

$$\frac{T_p}{T_e} = R_{low}\frac{1}{1+\beta^2} + R_{high}\frac{\beta^2}{1+\beta^2}, \tag{1}$$

where $\beta = \frac{P_{gas}}{P_{mag}}$ is the ratio of the gas pressure, P_{gas}, to the magnetic field pressure, $P_{mag} = B^2/2$, where B is the magnetic field strength. In the limit of $\beta \ll 1$, the temperature ratio asymptotically approaches $T_P/T_e \to R_{low}$, while in the limit of $\beta \gg 1$ the temperature ratio asymptotically approaches $T_P/T_e \to R_{high}$. We set R_{low} to 1 and we vary R_{high}. We varied the viewing angle, which is defined as the angle between the observer and the black hole spin axis. Finally, we overlaid our images with a circular mask and rotated them to change the PA, the projected angle between the image plane, and the black hole spin axis.

The GRMHD simulations were run up to $t_{final} = 10.000\,\mathcal{T}$ consisting of 1.000 snapshots with an interval of $10\,\mathcal{T}$. The correlation time of the image is $\sim 50\,\mathcal{T}$. In our data set, we used the last 100 snapshots of every spin value to capture the time evolution of the accretion flow. In these snapshots, the system is well evolved and the accretion flow has reached a quasi-steady state. We prevented our network from overfitting to single snapshots by randomly selecting ten snapshots as a validation set and training on the other snapshots. For each of these snapshots, we computed 200 images. Except for the spin, all parameters are

Table 1. Model parameters.

Parameter	Range
i	$[15°, 25°]$
$\dot{M}$	$[2 \times 10^{-6}\,M_\odot\,\mathrm{yr}^{-1}, 0.01\,M_\odot\,\mathrm{yr}^{-1}]$
R_{high}	$[1, 100]$
M	$[2 \times 10^9\,M_\odot, 8 \times 10^9\,M_\odot]$
PA	$[0°, 360°]$
a	$0, \pm 0.5, \pm 0.9375$

Notes. The parameter ranges used during data generation. The first five parameters are simulated continuously within the given range. The spin parameter a is only simulated at five values.

randomly picked from a uniform distribution between the parameter ranges given in Table 1. This ensures that there is no overlap between the training and validation sets. The mass prior is set such that it includes the one sigma range of the reported mass values of $3.5^{+0.9}_{-0.7} \times 10^9\,M_\odot$ and $6.6 \pm 0.4 \times 10^9\,M_\odot$ by Gebhardt et al. (2011) and Walsh et al. (2013).

Each parameter affects the image morphology differently. The viewing angle and PA influence the position of the jet and the asymmetry in the image. The density of the accretion flow determines the observed integrated flux and is related to the mass accretion rate $\dot{M}$. The size of the shadow is predominantly determined by the black hole mass M_{BH}. The value of R_{high} is related to the region the emission originates from, where a low value of R_{high} corresponds to a high concentration of emission in the disk and a high value corresponds to emission that predominantly originates from the jet. The black hole spin influences the geometry around the black hole and therefore the shape of the shadow and the asymmetry of the image. We show these effects in Sect. 3.1. A more detailed discussion of the effects of these parameters on the images can be found in Event Horizon Telescope Collaboration (2019e).

The viewing angle is sampled between 15° and 25° (Walker et al. 2018), the PA between 0° and 360°, R_{high} between 1 and 100, and M_{BH} between $2 \times 10^9\,M_\odot$ and $8 \times 10^9\,M_\odot$ (Gebhardt et al. 2011; Walsh et al. 2013). The mass accretion rate depends on the black hole mass, $\dot{M} \propto \mathcal{M}/M_{BH}$. To determine the prior of the mass accretion rate $\dot{M}$ for a fixed $M_{bh} = 6.5 \times 10^9\,M_\odot$, we manually fit to 1 Jy (Akiyama et al. 2015) at 230 GHz for $R_{high} = 1$ and $R_{high} = 100$ for every spin case at a mass of $M = 6.5 \times 10^9\,M_\odot$. The resulting range of priors is then extended with one order of magnitude to increase the scope of the training data and ensure the inclusion of the M 87* flux measurement. The resulting range of $\dot{M}$ is between 2×10^{-6} solar masses per year $(M_\odot\,\mathrm{yr}^{-1})$ and $0.01\,M_\odot\,\mathrm{yr}^{-1}$[3]. These parameter ranges are summarized in Table 1. We sampled all parameters linearly, except $\dot{M}$, which is sampled logarithmically because it covers a large range over multiple orders of magnitude. We predicted $\log(\dot{M})$ and converted this back to the original value after training in order to prevent from biasing our machine learning network. We applied a min-max normalization to all parameters. Finally, we convolved our images at 230 GHz with Gaussian beams of 5, 10, and $20\,\mu$as. The latter is the current nominal resolution of the EHT; future arrays might correspond to higher resolutions when either 345 GHz is added to the array or large baselines are realized. This is done as an approximation of a limited telescope resolution. In this work, we ignored other

---

[3] Or between $1.5 \times 10^{-8}\dot{M}_{Edd}$ and $7.3 \times 10^{-5}\dot{M}_{Edd}$, where $\dot{M}_{Edd}$ is the Eddington accretion rate for a black hole with mass $M = 6.5 \times 10^9\,M_\odot$.

telescope or measurement effects such as limited uv-coverage. The expected SVLBI resolution is sufficiently high that we do not convolve the images generated at 690 GHz; we assumed resolution as achieved by a set of telescopes in low Earth orbits, which are capable of obtaining resolutions of approximately $3\,\mu$as (for more information see Roelofs et al. 2019). As a result, we have five image libraries: one at 690 GHz and four at 230 GHz.

2.3. Neural networks

Recent technological developments led to advancements in the fields of deep learning and computer vision (Krizhevsky et al. 2012; Zeiler & Fergus 2014; Simonyan & Zisserman 2014; Szegedy et al. 2015; He et al. 2015). Computer vision using neural networks is typically done by training convolutional neural networks (CNNs; Lecun et al. 1998; Krizhevsky et al. 2012). We trained two CNNs in this work: a Bayesian regression neural network and a classification neural network.

Bayesian neural networks (BNNs) return a parameter estimation and a Bayesian motivated uncertainty (MacKay 1992; Gal 2016; Kendall & Gal 2017). A BNN predicts two types of uncertainties: the aleatoric and the epistemic uncertainty (Kiureghian & Ditlevsen 2009; Kendall & Gal 2017). The aleatoric uncertainty is associated with corruption in the data, for example due to a limited resolution, whereas the epistemic uncertainty is related to uncertainty in the model parameters, for example due to an insufficient amount of data or training. Although there are many types of uncertainties, they are generally categorized as either aleatoric or epistemic (Kiureghian & Ditlevsen 2009). Recent works have developed a fast and efficient method of approximating these uncertainties in machine learning (Gal & Ghahramani 2015a,b; Kendall & Gal 2017). A neural network is trained by optimizing a loss function. By choosing a Gaussian log-likelihood loss function, the network is also able to predict the aleatoric uncertainty. We split the final network layer, so it returns both a prediction and an aleatoric uncertainty. The epistemic uncertainty is obtained by using variational inference with a method called Monte Carlo dropout (MCD). In MCD, dropout layers (Srivastava et al. 2014) are included after every weighted layer in the network. These layers have a fixed probability, the dropout rate, to turn off individual neurons for a single forward pass of the data through the network. The dropout rate is tuned such that the fraction of validation examples that lay within a certain confidence interval correspond to those of a normal distribution (Perreault Levasseur et al. 2017). Dropout layers add a random component to the network, which results in repeated predictions of the same image giving varying outcomes. The collection of repeated predictions is used to sample the posterior probability distribution, and the variance of this distribution gives the epistemic uncertainty. The uncertainty can be combined by adding the epistemic variance to the mean aleatoric variance. By sampling the posterior with N predictions, the combined uncertainty σ is obtained by the following formula:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i\right)^2 + \frac{1}{N}\sum_{i=1}^{N}\sigma_{i,\mathrm{al}}^2, \quad (2)$$

where $\hat{y}$ is the network prediction and σ_{al} is the aleatoric uncertainty. For more details on this method, we refer to Kendall & Gal (2017) or Perreault Levasseur et al. (2017).

We used the regression BNN, hereafter network I, to predict the viewing angle i, mass accretion rate $\dot{M}$, plasma parameter R_{high}, black hole mass M_{BH} and t PA. We used the classifica-

tion network, hereafter network II, to predict the black hole spin a. We chose a classification network for this parameter because we only have five distinct values in the training sets. Network I is trained with a negative Gaussian log-likelihood loss function, described by

$$-L = \sum_i \frac{1}{2\hat{\sigma}_{i,\mathrm{al}}^2}\left|y_{n,i} - \hat{y}_{n,i}\right|^2 + \frac{1}{2}\log\hat{\sigma}_{i,\mathrm{al}}^2, \quad (3)$$

where $y_{n,k}$ is the true value of the k'th parameter in the n'th image. Network II predicts the black hole spin a. We train this parameter with a categorical cross-entropy loss function (Hastie et al. 2001), described by

$$L = -\sum_i^C y_{o,i}\log p_{o,i}, \quad (4)$$

where C is the number of classes, in our case five different spin values, $y_{o,i}$ is a binary indicator whether the class label i is the correct classification for observation o and $p_{o,i}$ is the predicted probability.

The architecture of networks I and II can be found in Fig. 1. Network I and II have a similar architecture, except for the output layers, where we differentiate between the classification network and the regression network. The latter network has a further differentiation between the network prediction and the aleatoric uncertainty prediction. We split our image libraries into a training set consisting of 90 000 images and a validation set consisting of the remaining 10 000 images to prevent the network from overfitting on features in the training set. In all but the output layers, we used a rectified linear unit (ReLU) activation function, described by $f(x) = \max(0, x)$ (Nair & Hinton 2010). We tuned the dropout rate as described in Sect. 2.3 and find 0.01 as our best value. The network is trained with the Adam optimizer (Kingma & Ba 2014) with Keras version 2.2.4 (Chollet 2015) and TensorFlow version 1.11.0 (Abadi et al. 2015). We set the initial learning rate on 0.001 and decreased it by a factor of 2 if the validation loss did not improve in two consecutive passes of the data through the network. We used a batch size of 32 during training. We used a random seed in TensorFlow of 1 to train our network, but we validated that our network can be trained independently of the chosen random seed.

3. Results

3.1. Image libraries

In Figs. 2–4, we show example images of the image libraries. Every image is generated by running RAPTOR with a unique set of parameters. The images show a central flux depression where the black hole is located, surrounded by a bright ring that coincides with the lensed emission ring (Falcke et al. 2000; Gralla et al. 2019; Johnson et al. 2020; Narayan et al. 2019). This ring scales with the black hole mass. Models with low values of R_{high} show extended emission features that originate from the accretion disk of the black holes. Many of the small-scale features are lost when the Gaussian beam is applied. With our network, we investigate what minimal resolution is required to make reliable parameter estimations by varying the Gaussian beam widths. We demand that false predictions are reflected by smaller network confidence through larger uncertainties on the prediction. These results can be seen in Sect. 3.2. The effective size of the EHT array is limited by the size of the Earth. Therefore, large improvements in the resolution require higher frequencies. Planned extensions of the EHT to
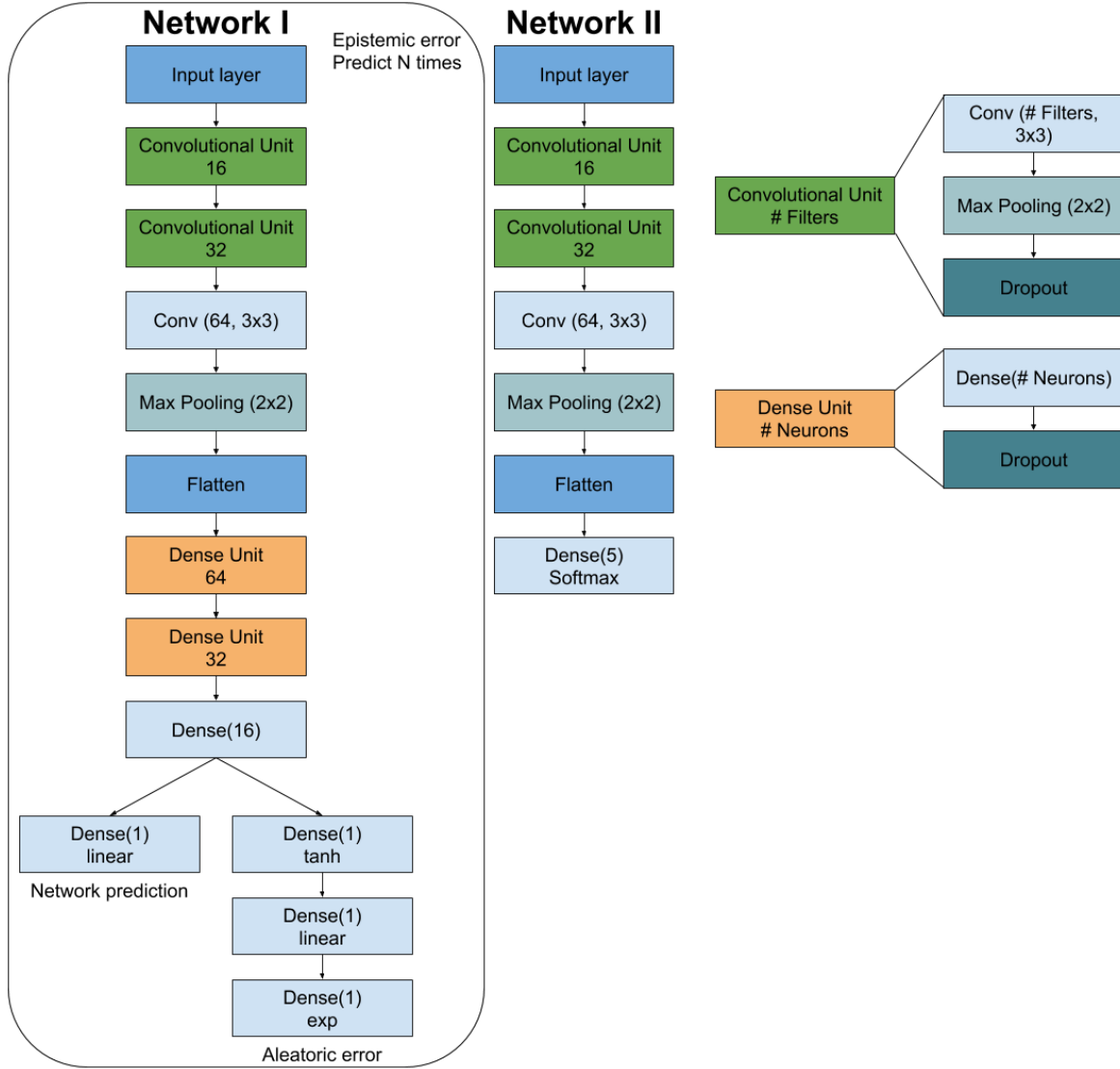
Fig. 1. Network architecture. After the flatten layer, network I branch out into a dense network per parameter, resulting in five unique network arms, which allow for parameter-specific learning. The next time the arms branch out into a network prediction and an aleatoric uncertainty prediction. To capture the epistemic error of network I, we make N predictions on the same image to sample the network posterior. Both networks have ReLU activation functions unless stated otherwise.

345 GHz would improve the resolution by ~40% (Event Horizon Telescope Collaboration 2019b). Further large improvements can be gained by switching to SVLBI. In Sect. 3.3, we show how our network performs at an SVLBI frequency of 690 GHz. The resolution of SVLBI experiments is expected to be sufficiently good enough to compare these experiments to the simulations without convolving them with a Gaussian beam.

3.2. Event Horizon Telescope

We plot the network I predictions of the 230 GHz image libraries as a function of the true values in Fig. 5. For each prediction, we determine the deviation for the correct prediction and weigh it with the uncertainty σ. This allows us to define how many points are predicted correctly within 1, 2, and 3σ, which corresponds to approximately 68%, 95%, and 99%. Without a Gaussian beam, the network can reliably predict the black hole parameters. A large scatter on the prediction indicates that the network is less confident, which is further reflected by a larger (mean) uncertainty. We show this uncertainty as a function of the deviation in Fig. 6. The horizontal line in this figure indicates the mean uncertainty. The values of the mean uncertainties can also be found in Table 2. With an increasing beam width, the scatter in the uncertainty and the mean value of the uncertainty increase.

In Fig. 7, we show a confusion matrix of the network II predictions. The neural network outputs a probability that an image belongs to a certain class and selects the class with the highest probability as network prediction. Without a Gaussian beam, the network has an accuracy of 95.9%, where most of the false classifications are either one spin value higher or lower than the correct value. The five classes are equally represented, and the entire validation set we show in this work contains 10 000 images. Up to a Gaussian beam of 10 μas, the network can discriminate between the different spin values, resulting in high accuracy and minor deviations between the different classes. At 20 μas beam width, this is no longer true, and this results in many more than half of the predictions being bad and larger discrepancies between the classes. We further investigate the predictions
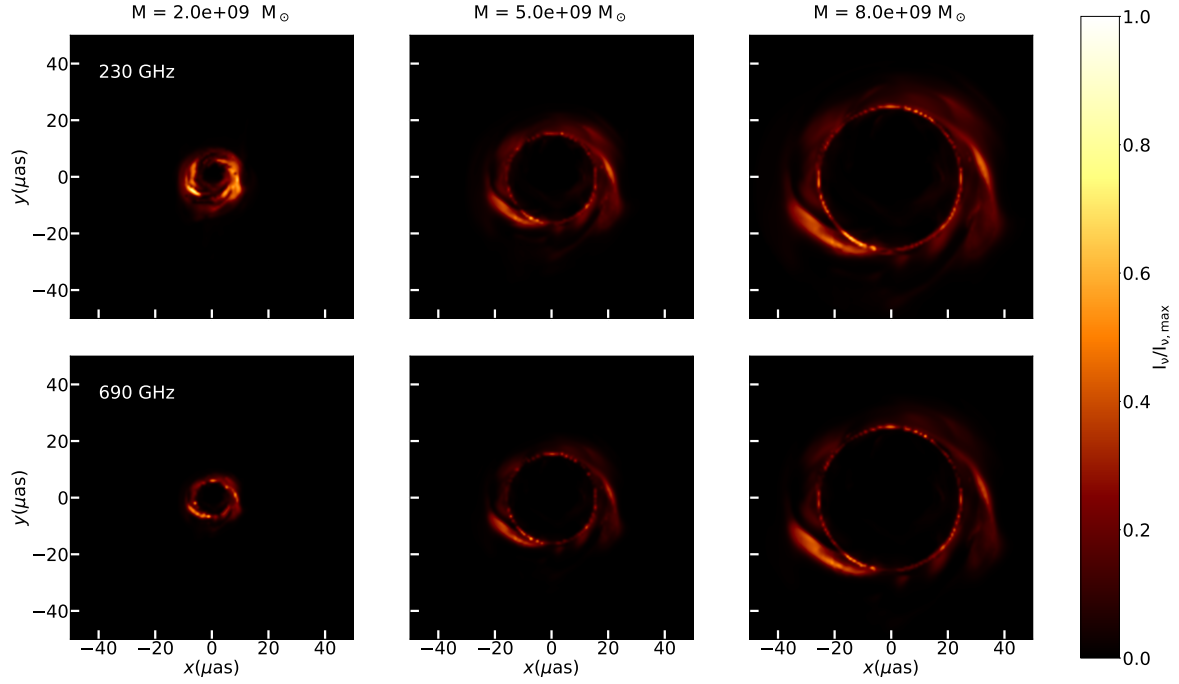
Fig. 2. Single snapshot synthetic images. *From left to right*: 2.0×10^9, 5.0×10^9, and 8.0×10^9 $M_\odot$. *From top to bottom*: 230 and 690 GHz. Images shown are representative images with a fixed flux of $F_{230\,\mathrm{GHz}} = 0.5$ Jy for model parameters $a = -1/2$, $i = 20°$, $R_{\mathrm{high}} = 50.0$, $\dot{M} = 10^{-4}$ $M_\odot$ yr^{-1}.
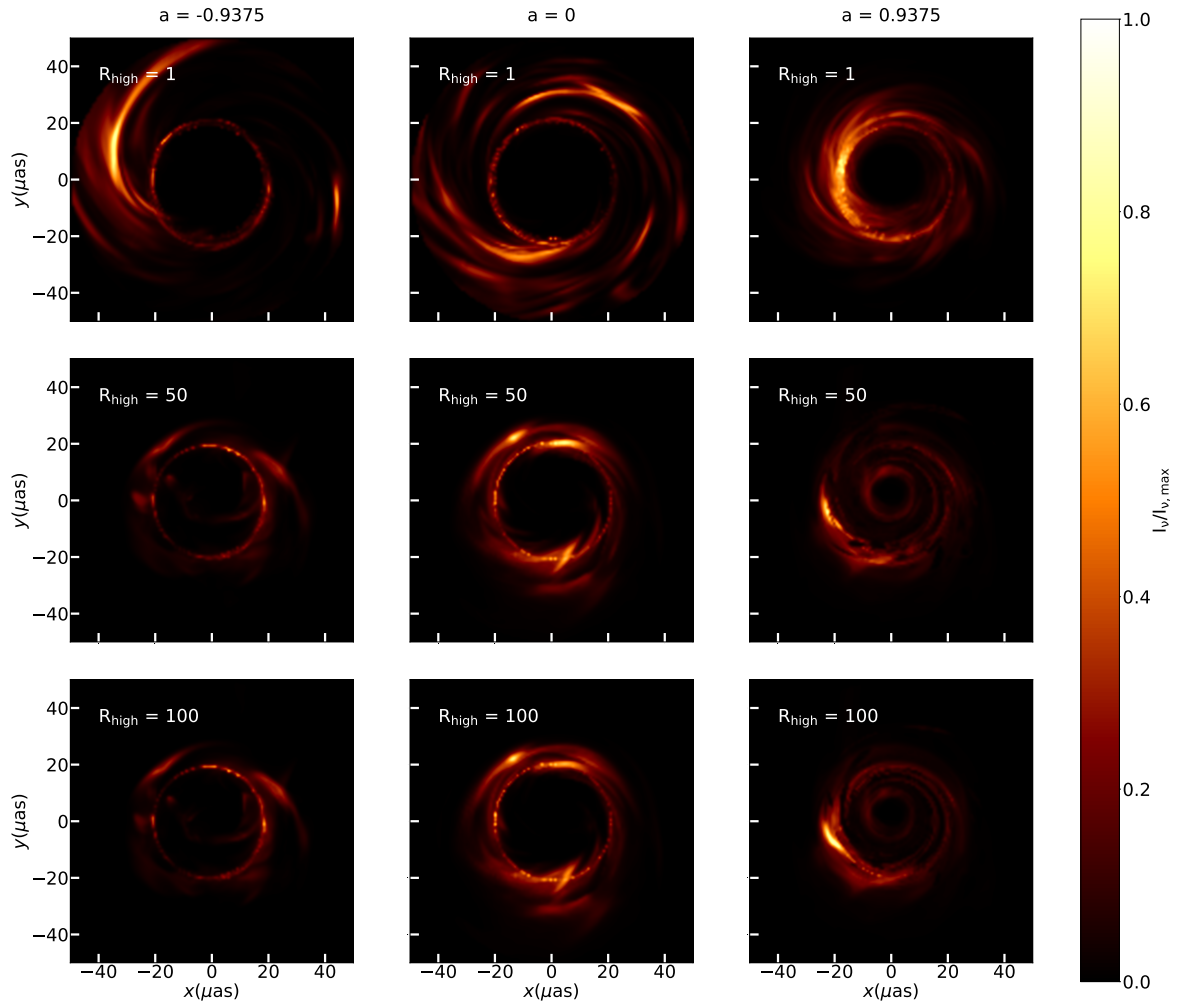
Fig. 3. Single snapshot synthetic images at 230 GHz. *From left to right*: spin values of -0.9375, 0, and 0.9375. *From top to bottom*: R_{high} values of 1, 50, and 100. Images shown are representative images with a fixed flux of $F_{230\,\mathrm{GHz}} = 0.5$ Jy for model parameters $i = 20°$ and $M = 6.5 \times 10^9$ $M_\odot$.
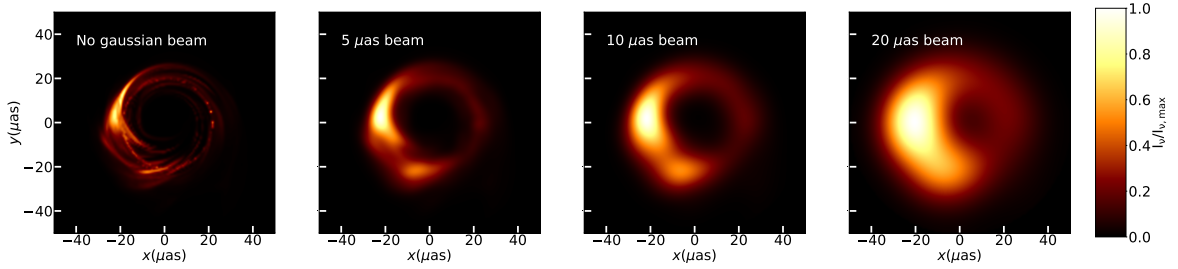
Fig. 4. Single snapshot synthetic images. *From left to right*: no Gaussian beam and Gaussian beam widths of 5, 10, and 20 μas. Model used is identical to the bottom right model in Fig. 3, with model parameters $a = 0.9375$, $R_{high} = 100$, $i = 20°$ and $M = 6.5 \times 10^9 \, M_{\odot}$.

by plotting the receiver operating characteristic (ROC) curves and the corresponding area under the curves (AUCs). In a ROC plot, the true positive rate (TPR) is plotted versus the false positive rate (FPR) as a function of the classification threshold. These quantities are given in Eq. (5), where TP stands for the number of true positives (prediction and truth are both true), P stands for the number of true values in the data set, and FN stands for the number of false negatives (prediction is false, while truth is true), as follows:

$$\text{TPR} = \text{TP/P} \quad \text{FPR} = \text{FN/P}. \tag{5}$$

The classification threshold is a minimal network probability that is required to belong to a certain class: with a threshold of zero, all spin values are compatible, whereas, with a threshold of one, only a perfect prediction is accepted. The AUC can be quantified with the integral defined as

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(x))\text{d}x. \tag{6}$$

In an accurate classifier, there is a high TPR at low FPR, which results in an AUC of 1. However, if the classifier cannot discriminate between the classes, they all are assigned approximately equal probabilities. Therefore, increasing the classification thresholds results in equally many true positives as false positives being accepted, which results in an AUC of 0.5. The ROC curves and the corresponding AUCs can be found in Fig. 8.

Up to a Gaussian beam width of 10 μas, Deep Horizon reliably predicts the parameters of this study. With larger beams, the images are more alike and, therefore, harder to distinguish. This results in larger (mean) uncertainties and lower AUCs. Some parameters are heavily affected by this (e.g., the viewing angle and R_{high}), whereas the parameters that predominantly affect large-scale features such as $\dot{M}$ (total flux) and M_{BH} (size of the shadow) are less affected by the Gaussian beam. For the viewing angle, i the addition of a 20 μas blur results in almost all predictions of the network to be close to the mean of the training data (for i at 20°). This is also reflected by the high uncertainty for all values, as can be seen in Fig. 6. The samples that result in network predictions, which are above the mean, show a majority of models with small black holes masses, high mass accretion rates, large R_{high} values, and larger viewing angles. These models have in common that they have a jet feature that potentially helps the network give a slight advantage to larger inclinations angles.

3.3. Space VLBI

We show the predictions of network I in Fig. 9, of network II in Fig. 10, and the mean uncertainty per parameter in Table 2. The performance at 690 GHz is comparable to 230 GHz without a Gaussian beam. The mass accretion rate, black hole mass,

and PA have a relatively low amount of scattering in the network predictions, which is reflected by the relatively low mean uncertainties. The average accuracy over all spin values of network II is 98.1%.

The expected data quality of SVLBI experiments allows for a direct comparison to mock data without convolving the images. We find that we can accurately recover all the parameters considered in this study. Therefore, future SVLBI missions would allow for more detailed measurements of SMBH systems.

4. Discussion

4.1. Network and data quality

In a machine learning algorithm, the training data are considered to be the ground truth. Therefore, the quality of the data set is important because any deviations of the simulation with respect to reality result in an unquantifiable uncertainty on the prediction. We have investigated the effects of increasing or decreasing the amount of training data and find that our algorithm is robust toward these changes. There are several ways we can further expand our image library to include more realistic mock observations. The first way is by including more parameters. In this proof of concept, we limit ourselves to SANE models for the accretion flow. However, the EHT observations are also in agreement with magnetically arrested disk (MAD; Bisnovatyi-Kogan & Ruzmaikin 1976; Narayan et al. 2003) models for the accretion flow. We also only investigate emission models based on a thermal distribution function for the electrons. Possible alternatives that relax the commonly used thermal distribution functions are either κ-distribution function (Davelaar et al. 2019) or power-law models (Dexter et al. 2012). Also, the choice of electron heating prescription limits the range of training data. Newly developed models include a secondary electron fluid that is evolved with the GRMHD simulation (Ressler et al. 2015; Ryan et al. 2018; Chael et al. 2019b), but these models highly depend on the choice for the underlying heating mechanism. In this proof of concept, we chose to use $R_{low} = 1$; this choice is identical to Event Horizon Telescope Collaboration (2019e). The value of R_{low} could, however, be smaller or lower than one, in general, the electron and proton fluid does not have to be in exact equilibrium in the jet sheath since electron heating in magnetized environments is shown to be more efficient (Howes 2010; Chandra et al. 2015; Rowan et al. 2017). To extend the predictive power of the network, a more diverse set of electron temperature models should be considered, for example, by including a larger range of R_{low} values. Another limitation of the training data can be found in the sampling of the spin parameter. Currently, only five spin simulations are present. The low sampling of the parameter space could result in overfitting on these examples when a classical regression
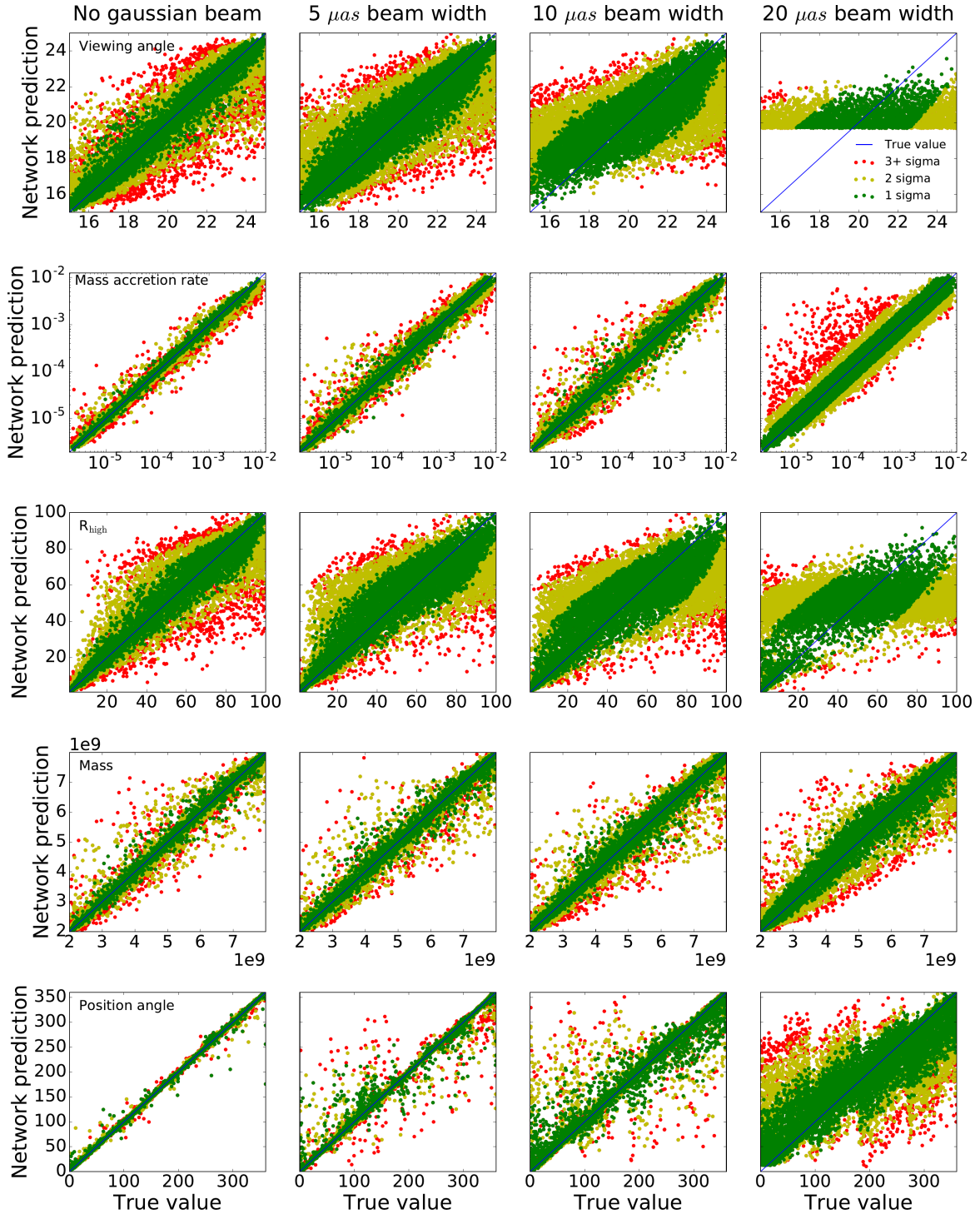
Fig. 5. Predictions of Network I at 230 GHz. The predictions of the 10 000 points in the validation set of network I at a frequency of 230 GHz with various amounts of Gaussian beam widths. The color coding is based on the deviation of the network prediction with respect to the true value, weighed by the network uncertainty. The red points indicate predictions that are correct within three or more σ. The units of the viewing angle and PA are degrees (°) of the mass accretion rate solar masses per year ($M_\odot$ yr^{-1}) and the mass is expressed in solar masses ($M_\odot$). *Left to right*: no Gaussian beam and Gaussian beam widths of 5, 10, and 20 μas. *Top to bottom*: viewing angle, mass accretion rate, R_{high}, mass, and PA. The sawtooth pattern in the PA at large Gaussian beams is discussed in Sect. 4.1.

network would be used. When applying our regression network to spin, large uncertainties are obtained owing to a limited amount of examples, which is expected for aleatoric uncertainties. To be able to perform regression on the spin, the number of values should be increased by at least a factor 2−4. Furthermore, we could also include images that are generated with theories

beyond GR. Thereby, the algorithm could learn to recognize if an image is compatible with GR or one of the alternatives (Mizuno et al. 2018). Including these additional parameters is beyond the scope of this proof of concept, but we would like to investigate this in future studies. Another method to improve our data sets is by including realistic telescope effects. In our data sets, we
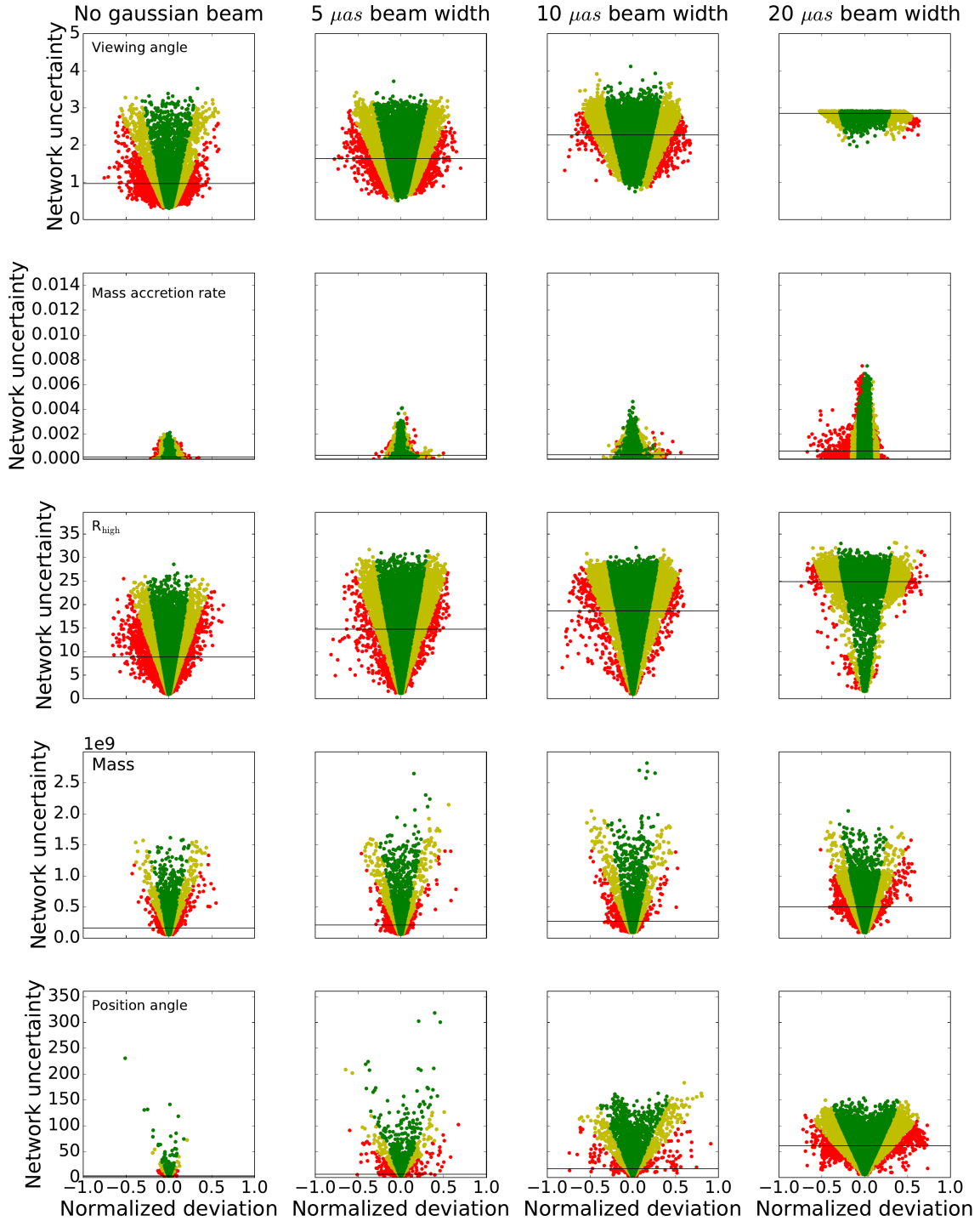
Fig. 6. Network I uncertainties at 230 GHz. The network uncertainties as a function of the deviation of the network prediction with respect to the true value. The deviation is normalized to the parameter ranges, where a deviation of one corresponds to a maximally wrong network prediction. The color coding, units, and order of the figure are similar to Fig. 5. The horizontal line indicates the mean predicted uncertainty.

approximate telescope resolution with a Gaussian beam but ignore effects such as thermal noise or telescope systematics. These effects are captured in SYMBA (Roelofs & Janssen, in prep.) and the eht-imaging (Chael et al. 2018, 2019a) Python package, which generate realistic synthetic data that can be reconstructed with imaging techniques. Improvements of the network could then go two ways, either by using reconstructed images based on synthetic data generated with GRMHD models or directly training the network on the visibility quantities

in the synthetic data sets; these are two options that should be compared in follow-up works. Finally, we also ignore various constraints on the measurements such as spectral energy distribution fitting, dynamical measurements, and polarization. Including this information in future studies could further improve our network efficiency.

Figure 5 shows asymmetries in the network predictions of the PA with a large Gaussian beam. Upon inspection of the data, we see that in many of the inaccurately predicted points, the jet

Table 2. Mean uncertainties of network I.

Beam width (μas)	230 GHz				690 GHz
	0	5	10	20	0
i (°)	1.0	1.6	2.3	2.9	1.0
$\dot{M}$ ($M_\odot$ yr^{-1})	1.8×10^{-4}	2.9×10^{-4}	3.6×10^{-4}	6.4×10^{-4}	2.4×10^{-4}
R_{high}	8.9	14.8	18.7	24.9	11.5
M_{BH} ($M_\odot$)	1.6×10^{8}	2.1×10^{8}	2.7×10^{8}	5.0×10^{8}	1.5×10^{8}
PA (°)	2.9	6.4	17.2	61.6	4.5

Notes. The individual predicted network uncertainties can be seen in Fig. 6. Left: 230 GHz. Right: same as left but at a frequency of 690 GHz.
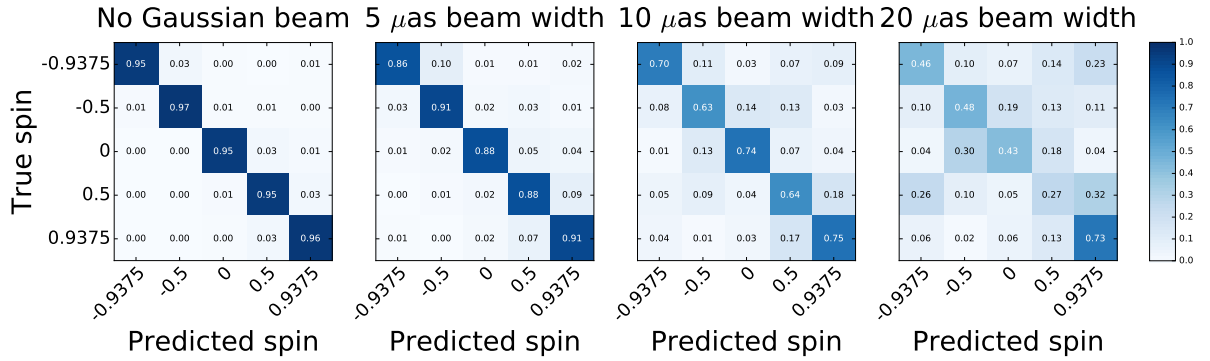
Fig. 7. Network II predictions at 230 GHz. The predictions of network II at a frequency of 230 GHz. The y-axis of the confusion matrix represents the true labels and the x-axis represents the network predictions. The individual cells show the accuracy of a given combination. The five classes are equally represented, and the total validation set contains 10.000 images.
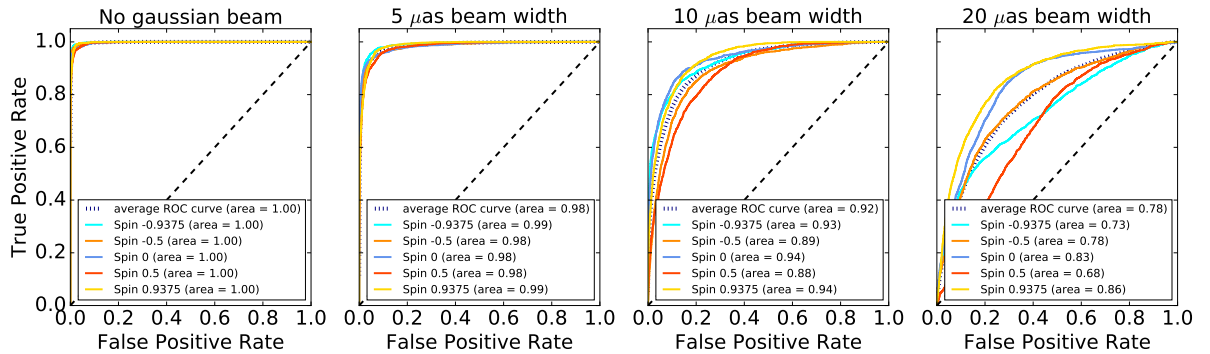

Fig. 8. Receiver operating characteristic curve and AUCs at 230 GHz. The ROC curves and corresponding AUCs at a frequency of 230 GHz. Accurate classifiers have an AUC of 1, whereas inaccurate classifiers have an AUC of 0.5. There are no large discrepancies between different spin values.

is not visible owing to the Gaussian beam. Instead, the Gaussian beam sometimes magnifies a local overabundance of plasma in the disk, which then shows up as if the jet is pointing in that direction. Furthermore, the loss function, as given in Eq. (3), does not capture the cyclicity of the parameter. Therefore, we introduce a bias in the PA that overestimates low PA values and underestimates high PA values. This causes the sawtooth pattern observed in Fig. 5. In future works, we want to investigate the effects of modifying the loss function to remove this asymmetry.

4.2. Time evolution

The environment near a black hole is a dynamic system. In our data generation, we use the last 100 snapshots of every spin value to capture these dynamics and prevent our network from overfitting on temporal features. We investigate the effects of the time evolution on our network by generating a new data set that has

a sufficiently large temporal separation to be used as an independent test. We find no large discrepancies between the independent test set and our standard data sets, and therefore, conclude that our network is not overfitting the temporal correlations within the data.

4.3. Comparison EHT

In Event Horizon Telescope Collaboration (2019f), three independent algorithms are employed to quantify the size, orientation, and shape of the asymmetric ring structure found in the 2017 EHT observations. These three methods are geometric crescent model fitting, GRMHD model fitting, and image domain feature extraction. In this paper, we present a fourth independent method that can be used for parameter estimations. In this subsection, we discuss the error budget obtained in the EHT methods. We focus on the measurement of the angular size
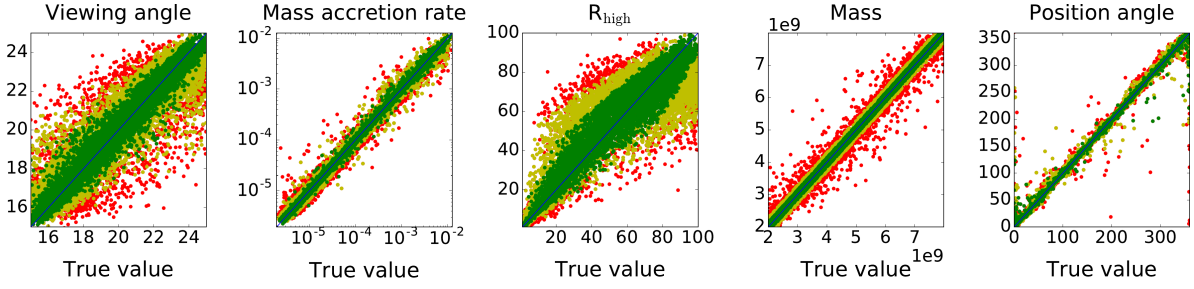
Fig. 9. Network I prediction at 690 GHz. The predictions of network I at a frequency of 690 GHz, no Gaussian beam applied. The color coding and units are similar to Fig. 5. *Left to right*: viewing angle, mass accretion rate, R_{high}, mass, and PA.
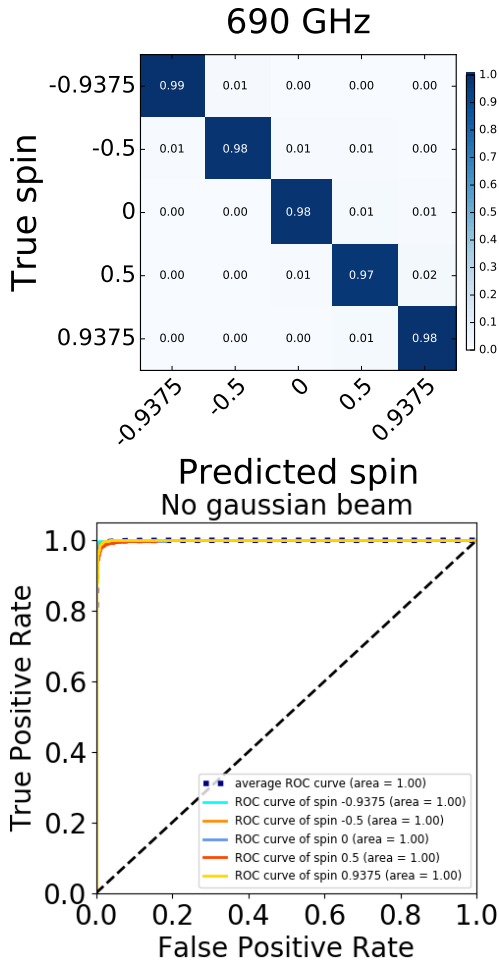

Fig. 10. Network II predictions at 690 GHz. *Top*: similar to Fig. 7 but at a frequency of 690 GHz. *Bottom*: similar to Fig. 8 but now for 690 GHz. For both panels, no Gaussian beam is applied.

corresponding to a gravitational radius, $\theta_g = R_g/D$, which is used to find the black hole mass by folding in a distance measurement for M 87.

The EHT reports three sources of uncertainty: a statistical uncertainty that corresponds to the width of the posterior, an observational uncertainty that corresponds to an incomplete (u, v) coverage, and unmodeled systematics and a theoretical uncertainty associated with the data being a single sample from a dynamic system. Of these three components, theoretical uncertainty is the largest component. Details on how the uncertainties are calculated per method are provided in the appendices of Event Horizon Telescope Collaboration (2019f). These indi-

vidual contributions are further classified as systematic uncertainties due to the GRMHD calibration of the method and a statistical uncertainty originating from the angular diameter measurement. The average values over the different methods after folding in the distance measurement are $\sigma_{sys} = 0.7 \times 10^9 \, M_\odot$ and $\sigma_{stat} = 0.2 \times 10^9 \, M_\odot$.

Although the same GRMHD simulations are used by the EHT and in this work, these uncertainties do not describe the same uncertainties as the uncertainties output by `Deep Horizon` because the methods to obtain these uncertainties are very different. Therefore, we cannot directly compare our network uncertainty to the uncertainty found within Event Horizon Telescope Collaboration (2019f). Furthermore, such a comparison would be beyond the scope of this paper because we do not test our network on the real image obtained by the EHT collaboration. However, we note that the mean uncertainties on the mass in Table 2 are of the same order of magnitude as the uncertainties found by the EHT. Although our method looks promising, further improvements and detailed method comparisons are required before we can apply this to observed data. One such comparison that requires further study is that, although our method differs from those described by the EHT, systematic correlations may remain as a result of the same underlying GRMHD simulations.

5. Conclusions

In this work, we present `Deep Horizon`, a combination of two convolutional deep neural networks that can recover input parameters of an image of an accreting SMBH. We create realistic mock observations and use these to show that our network can accurately recover the six parameters investigated in this study if we ignore limited telescope resolutions. We show that the current resolution of the Event Horizon Telescope is insufficient to determine all parameters of this study accurately, but is still sufficient to recover the mass and mass accretion rate accurately and could, therefore, confirm the results found by the Event Horizon Telescope collaboration. With future improvements to the resolution of images of black hole shadows, `Deep Horizon` would be able to recover more parameters. We investigated the case of space-based VLBI, which resulted in highly accurate parameter estimations.

Overall, the proof of concept presented in this paper shows that machine learning is an interesting parameter extracting tool for horizon scale observations that can be of great value for future tests of GR.

References

Abadi, M., Agarwal, A., Barham, P., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, Software Available from tensorflow.org

Akiyama, K., Lu, R.-S., Fish, V. L., et al. 2015, ApJ, 807, 150

Ball, N. M., Brunner, R. J., Myers, A. D., & Tcheng, D. 2006, ApJ, 650, 497

Bellinger, E. P., Angelou, G. C., Hekker, S., et al. 2016, ApJ, 830, 31

Bird, S., Harris, W. E., Blakeslee, J. P., & Flynn, C. 2010, A&A, 524, A71

Bisnovatyi-Kogan, G. S., & Ruzmaikin, A. A. 1976, Ap&SS, 42, 401

Bower, G. C., Goss, W. M., Falcke, H., Backer, D. C., & Lithwick, Y. 2006, ApJ, 648, L127

Broderick, A. E., Gold, R., Karami, M., et al. 2020, ApJ, submitted

Bronzwaer, T., Davelaar, J., Younsi, Z., et al. 2018, A&A, 613, A2

Cantiello, M., Blakeslee, J. P., Ferrarese, L., et al. 2018, ApJ, 856, 126

Chael, A., Bouman, K., Johnson, M., Blackburn, L., & Shiokawa, H. 2018, https://doi.org/10.5281/zenodo.1173414

Chael, A. A., Bouman, K. L., Johnson, M. D., et al. 2019a, Astrophysics Source Code Library [record ascl:1904.004]

Chael, A., Narayan, R., & Johnson, M. D. 2019b, MNRAS, 486, 2873

Chandra, M., Gammie, C. F., Foucart, F., & Quataert, E. 2015, ApJ, 810, 162

Chollet, F. 2015, Keras, https://keras.io

Davelaar, J., Mościbrodzka, M., Bronzwaer, T., & Falcke, H. 2018, A&A, 612, A34

Davelaar, J., Olivares, H., Porth, O., et al. 2019, A&A, 632, A2

Dexter, J., McKinney, J. C., & Agol, E. 2012, MNRAS, 421, 1517

Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019a, ApJ, 875, L1

Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019b, ApJ, 875, L2

Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019c, ApJ, 875, L3

Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019d, ApJ, 875, L4

Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019e, ApJ, 875, L5

Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019f, ApJ, 875, L6

Fadely, R., Hogg, D. W., & Willman, B. 2012, ApJ, 760, 15

Falcke, H., Melia, F., & Agol, E. 2000, ApJ, 528, L13

Fan, X., Li, J., Li, X., Zhong, Y., & Cao, J. 2019, Sci. China Phys. Mech. Astron., 62, 969512

Fish, V. L., Shea, M., & Akiyama, K. 2020, Adv. Space Res., 65, 821

Fishbone, L. G., & Moncrief, V. 1976, ApJ, 207, 962

Fromm, C. M., Younsi, Z., Baczko, A., et al. 2019, A&A, 629, A4

Gal, Y. 2016, PhD Thesis, University of Cambridge

Gal, Y., & Ghahramani, Z. 2015a, ArXiv e-prints [arXiv:1506.02142]

Gal, Y., & Ghahramani, Z. 2015b, ArXiv e-prints [arXiv:1506.02158]

Gebhardt, K., Adams, J., Richstone, D., et al. 2011, ApJ, 729, 119

George, D., & Huerta, E. A. 2018, Phys. Lett. B, 778, 64

Goddi, C., Falcke, H., Kramer, M., et al. 2017, Int. J. Mod. Phys. D, 26, 1730001

Goodman, J., & Narayan, R. 1989, MNRAS, 238, 995

Gralla, S. E., Holz, D. E., & Wald, R. M. 2019, Phys. Rev. D, 100, 024018

Hastie, T., Tibshirani, R., & Friedman, J. 2001, The Elements of Statistical Learning, Springer Series in Statistics (New York, NY, USA: Springer New York Inc.)

He, K., Zhang, X., Ren, S., & Sun, J. 2015, ArXiv e-prints [arXiv:1512.03385]

Hendriks, L., & Aerts, C. 2019, PASP, 131, 108001

Hezaveh, Y. D., Perreault Levasseur, L., & Marshall, P. J. 2017, Nature, 548, 555

Hon, M., Stello, D., & Yu, J. 2017, MNRAS, 469, 4578

Howes, G. G. 2010, MNRAS, 409, L104

Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90

Jacobs, C., Glazebrook, K., Collett, T., More, A., & McCarthy, C. 2017, MNRAS, 471, 167

Johannsen, T., & Psaltis, D. 2010, ApJ, 718, 446

Johnson, M. D., & Gwinn, C. R. 2015, ApJ, 805, 180

Johnson, M. D., Lupsasca, A., Strominger, A., et al. 2020, Sci. Adv., 6, eaaz1310

Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python [Online]

Kendall, A., & Gal, Y. 2017, in Advances in Neural Information Processing Systems 30, eds. I. Guyon, U. V. Luxburg, S. Bengio, et al. (Curran Associates, Inc.), 5574

Kerr, R. P. 1963, Phys. Rev. Lett., 11, 237

Kim, E. J., & Brunner, R. J. 2017, MNRAS, 464, 4463

Kim, E. J., Brunner, R. J., & Carrasco Kind, M. 2015, MNRAS, 453, 507

Kingma, D. P., & Ba, J. 2014, ArXiv e-prints [arXiv:1412.6980]

Kiureghian, A. D., & Ditlevsen, O. 2009, Struct. Saf., 31, 105

Krizhevsky, A., Sutskever, I., & Hinton, G. E. 2012, in Advances in Neural Information Processing Systems 25, eds. F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Curran Associates, Inc.), 1097

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. 1998, Proc. IEEE, 2278

Lukic, V., & Brüggen, M. 2017, in Astroinformatics, eds. M. Brescia, S. G. Djorgovski, E. D. Feigelson, G. Longo, & S. Cavuoti, IAU Symp., 325, 217

MacKay, D. J. C. 1992, Neural Comput., 4, 448

Millman, K. J., & Aivazis, M. 2011, Comput. Sci. Eng., 13, 9

Mizuno, Y., Younsi, Z., Fromm, C. M., et al. 2018, Nat. Astron., 2, 585

Mościbrodzka, M., Falcke, H., & Shiokawa, H. 2016, A&A, 586, A38

Mościbrodzka, M., Dexter, J., Davelaar, J., & Falcke, H. 2017, MNRAS, 468, 2214

Nair, V., & Hinton, G. E. 2010, in Rectified Linear Units Improve Restricted Boltzmann Machines, eds. J. Fürnkranz, & T. Joachims (Omnipress), 807

Narayan, R., & Goodman, J. 1989, MNRAS, 238, 963

Narayan, R., Igumenshchev, I. V., & Abramowicz, M. A. 2003, PASJ, 55, L69

Narayan, R., SÄdowski, A., Penna, R. F., & Kulkarni, A. K. 2012, MNRAS, 426, 3241

Narayan, R., Johnson, M. D., & Gammie, C. F. 2019, ApJ, 885, L33

Odewahn, S. C., Stockwell, E. B., Pennington, R. L., Humphreys, R. M., & Zumach, W. A. 1992, in Digitised Optical Sky Surveys, eds. H. T. MacGillivray, & E. B. Thomson, Astrophys. Space Sci. Lib., 174, 215

Oliphant, T. E. 2007, Comput. Sci. Eng., 9, 10

Olivares, H., Porth, O., Davelaar, J., et al. 2019, A&A, 629, A61

Palumbo, D., Johnson, M., Doeleman, S., Chael, A., & Bouman, K. 2018, Am. Astron. Soc. Meet. Abstr., 231, 347.21

Perreault Levasseur, L., Hezaveh, Y. D., & Wechsler, R. H. 2017, ApJ, 850, L7

Petrillo, C. E., Tortora, C., Chatterjee, S., et al. 2017, MNRAS, 472, 1129

Porth, O., Olivares, H., Mizuno, Y., et al. 2017, Comput. Astrophys. Cosmol., 4, 1

Porth, O., Chatterjee, K., Narayan, R., et al. 2019, ApJS, 243, 26

Psaltis, D., Özel, F., Chan, C.-K., & Marrone, D. P. 2015, ApJ, 814, 115

Ressler, S. M., Tchekhovskoy, A., Quataert, E., Chandra, M., & Gammie, C. F. 2015, MNRAS, 454, 1848

Roelofs, F., Falcke, H., Brinkerink, C., et al. 2019, A&A, 625, A124

Rowan, M. E., Sironi, L., & Narayan, R. 2017, ApJ, 850, 29

Ryan, B. R., Ressler, S. M., Dolence, J. C., Gammie, C., & Quataert, E. 2018, ApJ, 864, 126

Schwarzschild, K. 1916, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, Phys.-Math. Klasse, 189

Sevilla-Noarbe, I., & Etayo-Sotos, P. 2015, Astron. Comput., 11, 64

Shen, H., Huerta, E. A., & Zhao, Z. 2019, ArXiv e-prints [arXiv:1903.01998]

Simonyan, K., & Zisserman, A. 2014, ArXiv e-prints [arXiv:1409.1556]

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, J. Mach. Learn. Res., 15, 1929

Suchkov, A. A., Hanisch, R. J., & Margon, B. 2005, AJ, 130, 2439

Szegedy, C., Liu, W., Jia, Y., et al. 2015, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1

Tchekhovskoy, A., Narayan, R., & McKinney, J. C. 2011, MNRAS, 418, L79

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Comput. Sci. Eng., 13, 22

Vasconcellos, E. C., de Carvalho, R. R., Gal, R. R., et al. 2011, AJ, 141, 189

Walker, R. C., Hardee, P. E., Davies, F. B., Ly, C., & Junor, W. 2018, ApJ, 855, 128

Walsh, J. L., Barth, A. J., Ho, L. C., & Sarzi, M. 2013, ApJ, 770, 86

Weir, N., Fayyad, U. M., Djorgovski, S. G., & Roden, J. 1995, PASP, 107, 1243

Zeiler, M. D., & Fergus, R. 2014, European Conference on Computer Vision (Springer), 818