High-contrast spectroscopy of exoplanet atmospheres
Landman, R.

Citation
Landman, R. (2024, June 11). *High-contrast spectroscopy of exoplanet atmospheres*. Retrieved from https://hdl.handle.net/1887/3762663

6 | Making the unmodulated Pyramid wavefront sensor smart.

Closed-loop demonstration of neural network wavefront reconstruction with MagAO-X

Almost all current and future high-contrast imaging instruments will use a Pyramid wavefront sensor (PWFS) as a primary or secondary wavefront sensor. The main issue with the PWFS is its nonlinear response to large phase aberrations, especially under strong atmospheric turbulence. Most instruments try to increase its linearity range by using dynamic modulation, but this leads to decreased sensitivity, most prominently for low-order modes, and makes it blind to petal-piston modes. In the push toward high-contrast imaging of fainter stars and deeper contrasts, there is a strong interest in using the PWFS in its unmodulated form. Here, we present closed-loop lab results of a nonlinear reconstructor for the unmodulated PWFS of the Magellan Adaptive Optics eXtreme (MagAO-X) system based on convolutional neural networks (CNNs). We show that our nonlinear reconstructor has a dynamic range of >600 nm root-mean-square (RMS), significantly outperforming the linear reconstructor that only has a 50 nm RMS dynamic range. The reconstructor behaves well in closed loop and can obtain >80% Strehl at 875 nm under a large variety of conditions and reaches higher Strehl ratios than the linear reconstructor under all simulated conditions. The CNN reconstructor also achieves the theoretical sensitivity limit of a PWFS, showing that it does not lose its sensitivity in exchange for dynamic range. The current CNN's computational time is 690 μs, which enables loop speeds of >1 kHz. On-sky tests are foreseen soon and will be important for pushing future high-contrast imaging instruments toward their limits.

6.1   Introduction

The direct imaging of extrasolar planets and circumstellar disks requires high contrast at small angular separations (Bowler, 2016). A crucial component of ground-based high-contrast imaging (HCI) instruments is an extreme adaptive optics (AO) system that corrects for optical distortions from the atmosphere and the instrument itself (Guyon, 2018). To push HCI toward fainter stars and deeper contrasts, AO systems need highly sensitive wavefront sensors (WFSs) that optimally use all the photons to sense the incoming wavefront. Most current and future HCI instruments will use or are already using a Pyramid wavefront sensor (PWFS; Ragazzoni, 1996) as their main WFS (e.g., Pinna et al., 2016; Kasper et al., 2021; Males et al., 2022a,b; Haffert et al., 2022; Bond et al., 2022). Furthermore, upgrades of other HCI instruments are also planning to use a PWFS as a primary or second-stage WFS (Fitzsimmons et al., 2020; Perera et al., 2022; Boccaletti et al., 2022). The increased preference for the PWFS over the Shack-Hartmann WFS is mostly due to its improved sensitivity (Ragazzoni & Farinato, 1999; Chambouleyron et al., 2023). This enhanced sensitivity allows AO systems to run at faster loop speeds, improving the achievable contrast at small angular separations. Additionally, its sensitivity can further be optimized by changing the binning fraction of the detector.

The main issue with the unmodulated PWFS is its nonlinear response to incoming wavefront aberrations, especially in the presence of strong turbulence (Esposito & Riccardi, 2001). Without modulation, the linearity range of the PWFS is much smaller than the phase aberrations that are typically observed from atmospheric turbulence. As a result, most instruments use dynamic modulation to increase its linearity. However, this modulation reduces its sensitivity, especially to low-order modes, decreasing its performance on faint stars and limiting its loop speed. Additionally, it strongly reduces its ability to sense petal-piston modes, which are crucial for upcoming giant segmented telescopes (Bertrou-Cantou et al., 2022; Hedglen et al., 2022; Engler et al., 2022). Even when modulated, the PWFS still exhibits a nonlinear response to large aberrations. Furthermore, if a deformable mirror (DM) upstream of the PWFS is used to dig a dark hole or correct non-common path aberrations, the PWFS may need to operate with a nonzero offset, limiting its linearity range even more.

Many different solutions have been suggested to mitigate the nonlinearity of the PWFS. Some works have proposed a first-order correction to this nonlinearity in the form of an optical gain compensation (Deo et al., 2019; Chambouleyron et al., 2020, 2021). However, this requires knowledge of the turbulence statistics and cannot account for nonlinear modal cross-talk. On the other hand, nonlinear reconstructors provide a software-based solution that can do higher-order cor-

6

rection without a reduction in sensitivity. Nonlinear reconstructors can be split into model-based and data-driven algorithms. Model-based methods rely on an accurate model of the optical system of the PWFS and accurate WFS to DM calibration. Model-based methods include reconstructors based on gradient descent (Hutterer & Ramlau, 2018; Frazin, 2018; Hutterer et al., 2023) and the Gerchberg-Saxton algorithm (Chambouleyron et al., 2024). These methods often need multiple iterations to converge to an accurate wavefront estimation. Alternatively, data-driven nonlinear reconstructors have shown promising results. These methods use the ability of neural networks (NNs) to approximate arbitrary functions to learn the inverse relation between WFS measurements and the incoming wavefront (Swanson et al., 2018; Landman & Haffert, 2020; Archinuk et al., 2023; Wong et al., 2023). While these methods show promising results, there is a lack of closed-loop lab and on-sky demonstrations of these techniques. Finally, nonlinear or adaptive controllers may help alleviate some of the nonlinearity problems of the PWFS using time-domain information (Landman et al., 2021; Wong et al., 2021; Deo et al., 2021; Haffert et al., 2021; Nousiainen et al., 2022; Pou et al., 2022).

In this work we present lab-based closed-loop tests of a convolutional neural network (CNN) reconstructor for the unmodulated PWFS of the Magellan Adaptive Optics eXtreme (MagAO-X) system. Section 6.2 discusses the calibration and architecture of the CNN reconstructor. Section 6.3 presents the open-loop performance of the reconstructor compared to a linear reconstructor, while Section 6.4 compares the performance in a closed loop. Finally, Section 6.5 summarizes the results and lists our conclusions.

6.2   Methods

6.2.1   MagAO-X

MagAO-X is an extreme AO instrument for the 6.5 meter Magellan Clay Telescope at Las Campanas Observatory (LCO) in Chile. The instrument is shipped back and forth between Steward Observatory in Tucson, Arizona, and LCO. Shipping the instrument back and forth allows us 100% access to the instrument when it is in Tucson. All the experiments shown in this work were performed at Steward Observatory.

MagAO-X is split into two optical benches that are connected by a periscope relay. The upper bench has the telescope simulator that is fed by a single-mode fiber-coupled super continuum laser (SuperK from NKT photonics). This source goes through the telescope simulator that generates an f/11 beam, which is equal
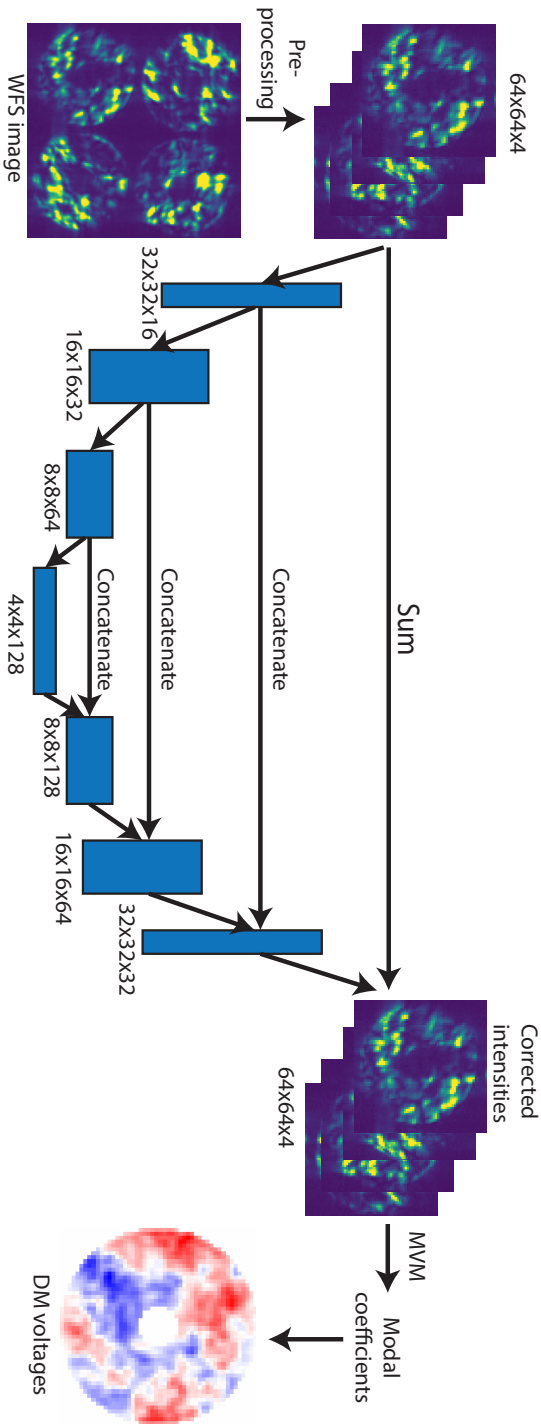
Figure 6.1: Visualization of the reconstruction pipeline and our NN architecture. The model consists of a U-net architecture with a nonlinear encoder and a decoder with skip connections. The nonlinear part uses leaky ReLU activation functions. Connections in the encoder part of the network use convolutional layers with a stride of 2, while the decoder part consists of transpose convolutional layers to ensure matching dimensions.

to focal ratio of the Magellan Telescope. A pupil mask is used in an intermediate pupil to create the exact aperture of the Magellan Telescope. This beam is injected into the instrument and passes through several optics on the upper optical bench. The most important ones are the DMs. MagAO-X uses a woofer-tweeter architecture with a ALPAO-97 DM as woofer and a Boston Micromachines 2K tweeter (Males et al., 2018; Close et al., 2018; Males et al., 2022a). The beam is relayed by the periscope mirrors to the lower optical table, which contains the PWFS and the science instrumentation. A beamsplitter directly after the periscope system splits the light off into two paths; one for the PWFS for wavefront sensing and one for the science instrumentation. The AO beam goes through some additional flat mirrors and is collimated onto a high-speed piezo modulator. This beam is focused on the MagAO-X pyramid prism (Schatz et al., 2018). A custom triplet camera lens collimates the beam onto an electron-multiplying CDD (OCAM2K camera). The four PWFS pupils are sampled by 56 pixels across the pupil with a separation of 60 pixels. The science beam is focused by an off-axis parabola that creates a f/69 beam onto the science cameras. The science cameras sample the point spread function (PSF) with 3 pixels per λ/D at Hα, which is 5.98 mas pixel^{-1} on-sky (Long et. al. in prep).

6.2.2 Training data

One of the most important components toward obtaining a data-driven nonlinear reconstructor is generating an appropriate training set. This training set needs to be representative of the data that are seen on-sky. This is not trivial, as the distribution of aberrations in closed loop depends on among others on the turbulence statistics, loop gain and noise propagation for different spatial frequencies. To perform well over a large variety of conditions in a closed loop, we generate phase screens with random power-law power spectral densities and root-mean-square (RMS) phase. We uniformly sampled power-law indices between -1 and -3 and RMS values uniformly distributed in log-space between 0.2 nm and 600 nm. This ensures that the dataset contains the full range of aberrations that are seen on-sky in a closed-loop setting, from the large aberrations when closing the loop to the small aberrations after convergence. These phase screens were then projected on the controllable modes of the tweeter. For the reconstruction, we chose to reconstruct 1000 modes in a Fourier basis. We only reconstructed and controlled 1000 modes in to allow for the simulated turbulence, which is introduced using the same DM, to include a larger number of modes than we control. This mimics the situation on-sky, where we have to deal with the fitting error and subsequent optical gain effects. The modal coefficients for the training set were obtained by projecting the phase screens onto this modal basis. In total, we generated 100,000

phase screens, which were used for training, validation, and testing in a 60%, 15%, and 25% split. All the data were collected with the ND2 neutral density filter inserted, which reduces the flux level by a factor of 100. This creates a photon flux that is roughly equivalent to observing a $\sim$ 0th magnitude star (I-band) at 2 kHz with MagAO-X.

6.2.3 Neural network

Artificial NNs are arbitrary function approximators. They consist of multiple sequential matrix vector multiplications (MVMs) with a nonlinear activation function in between, enabling them to learn nonlinear input-output relations. To make the model fitting tractable one can use domain knowledge to limit the amount of free parameters. The most prominent example of this are CNNs, which use the assumption that features are local and translationally invariant to drastically reduce the amount of free parameters. These CNNs have been used extensively for image processing tasks over the last decade and have revolutionized many fields of science. Since we are also dealing with images here, we chose to use CNNs as our model architecture. The architecture used here closely follows the one used in Landman et al. (2022) and is based on U-net (Ronneberger et al., 2015), which uses skip connections to reduce the vanishing gradient problem. The input consists of square images of 64x64 pixels around the center of the pyramid pupils concatenated along its depth, resulting in an input of 64x64x4, as visualized in Fig. 6.1. These images were normalized by their total intensity and scaled by a fixed constant such that the input data are between 0 and 1. This passes trough the nonlinear part of the network, which consists of an encoder and decoder. The leaky ReLU activation was used for each of the layers in the nonlinear part of the model. The output of this nonlinear part is a correction on the intensity image, which is added to the original image. This is subsequently mapped to the modal coefficients with a standard MVM, or a "fully connected" layer. These modal coefficients were scaled by a fixed constant, given by the standard deviation in the training set. The MVM here is optimized as part of the NN and is not the same as the linear reconstructor derived in Section 6.2.4. The main advantage of this architecture is the presence of a direct linear connection between the input image and the output modal coefficients, allowing it to mimic a linear reconstructor with ease. This integrates the linear term that was found necessary for stable closed-loop operation in Landman & Haffert (2020) and Pou et al. (2022). The loss function J used here is based on the relative loss used in Landman & Haffert (2020) and is given by the ratio of the residual RMS error divided by the input RMS, with a constant ϵ to avoid divergence for very small input RMS. The loss is defined as follows:

$$J = \left\langle \frac{\sqrt{\sum_i (y_{\text{true},i} - y_{\text{pred},i})^2}}{\sqrt{\sum_i y_{\text{true},i}^2} + \epsilon} \right\rangle, \tag{6.1}$$

where <> denotes the mean over a sampled batch, $y_{\text{true},i}$ the applied modal coefficients for mode i and $y_{\text{pred},i}$ the predicted modal coefficients by the NN for that mode. This loss ensures that the model focuses in equal amounts on the small and large aberrations. If we simply used the residual RMS as the loss, this would disproportionately focus on the large aberrations. This would lead to suboptimal closed-loop performance, as the residuals in closed loop are often small (Landman & Haffert, 2020). Throughout this work we use $\epsilon = 2$ nm, as this was found to be the limiting precision for the linear model.

We found that this model was initially overfitting to the training data and we subsequently added regularization. We used Dropout (Srivastava et al., 2014) layers after the first two convolutional layers with 10% dropout and one with 30% dropout before the MVM. Additionally, we used L_2 regularization with a value of 10^{-5} in the final MVM layer, as that layer has most of the free parameters. The Adam optimizer (Kingma & Ba, 2014) was used to train the CNN with an initial learning rate of 0.003, which was decayed after every epoch by a factor of 0.96, and a batch size of 32. Training of the model took about 1 hour on a single GeForce RTX 2080 Ti GPU.

6.2.4 Linear model

The linear model was constructed using the same data that were used to train the CNN. However, we only included data with an RMS phase smaller than 20 nm to ensure we are in a regime where the WFS is mostly linear. We calculated the reduced PWFS intensities by subtracting the zero-point reference and dividing by the total intensity. The reconstruction matrix, R, was then obtained using a regularized linear least squares:

$$R = (X^T X + \rho \mathbb{I})^{-1} X^T Y, \tag{6.2}$$

where X is the matrix with in its rows the measured reduced intensities, Y the matrix with in its rows the applied modal coefficients, ρ the regularization parameter and $\mathbb{I}$ the identity matrix. We only included the first 14,000 measurements in the regression to make the inversion feasible, which was calculated using a singular value decomposition. The regularization parameter was optimized on a test dataset of 3,000 samples. This was done by trying different values of ρ and choosing the one that gave the lowest reconstruction RMS on the test dataset. We found an optimal value of $\rho = 3 \times 10^{-4}$ for the regularization parameter.
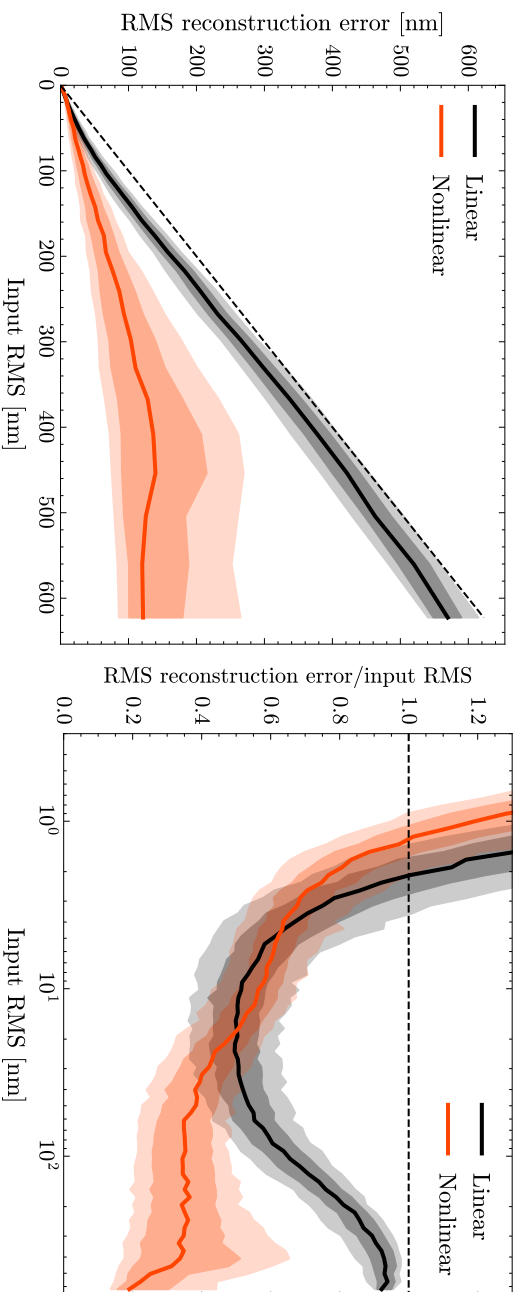
6

Figure 6.2: Comparison of the reconstruction accuracy of the nonlinear CNN with a linear model on the test dataset. The solid line indicates the mean value, and the colored region shows the 68% and 95% confidence intervals. The dotted line indicates where the RMS error of the reconstruction is equal to the input RMS, i.e., where there is no improvement. Left: Residual RMS error of the prediction as a function of the RMS of the input wavefront, shown on a linear scale. Right: Improvement factor, defined as the ratio between the residual RMS and the input RMS, as a function of the input RMS, which is shown on a logarithmic scale.

6

6.3   Open-loop reconstruction

6.3.1   Reconstruction accuracy

First, we compared the ability of the models to reconstruct the wavefront in an open-loop setting. To do this, we used the last 25% of the generated data and evaluated the RMS error of the reconstruction, defined as

$$\text{RMS} = \sqrt{\sum_i (y_{\text{true},i} - y_{\text{pred},i})^2}. \tag{6.3}$$

The reconstruction RMS error as a function of the RMS of the input wavefront is shown in Fig. 6.2. We see that the NN has significantly reduced reconstruction RMS compared to the linear model. For the largest aberrations in our dataset (600 nm) there is a factor of 3 to 4 improvement in reconstruction accuracy with respect to the linear model, showing the increased dynamic range that can be obtained with a nonlinear reconstructor. The improvement factor, which is given by the residual RMS after reconstruction divided by the RMS of the input wavefront is shown on the right side of Fig. 6.2. When this curve intersects one the reconstruction does not lead to an improvement in wavefront quality. The intersection at the smallest input RMS therefore represents the sensitivity limit, while the intersection at the largest RMS determines the dynamic range. We observe that the NN is able to obtain a better sensitivity limit than the linear model. This is likely already due to nonlinearities, as the linear model was calibrated with wavefronts until 20 nm RMS, which is outside the fully linear regime of the unmodulated PWFS. The sensitivity limit for both reconstructors will be studied in more detail in the next section.

6.3.2   Reconstruction accuracy for fainter stars

The main advantage of the unmodulated PWFS is its increased sensitivity, which allows for better AO performance on faint stars (Guyon, 2005; Agapito et al., 2023). However, this performance benefit requires that the nonlinear reconstructor does not have stronger noise amplification than the linear model. Furthermore, in the case of noisy measurements, the NN may struggle to resolve nonlinear structure in the data, as this can be washed out by noise. To test the performance of the reconstructor on noisy data, we artificially increased the noise level of the test data and evaluated its open-loop reconstruction performance. This was done by adding Poisson noise for a given stellar magnitude. We ignored read noise as photon noise dominates for most WFSs using modern electron-multiplying CCDs, especially for the relatively bright natural guide stars in HCI. We converted the
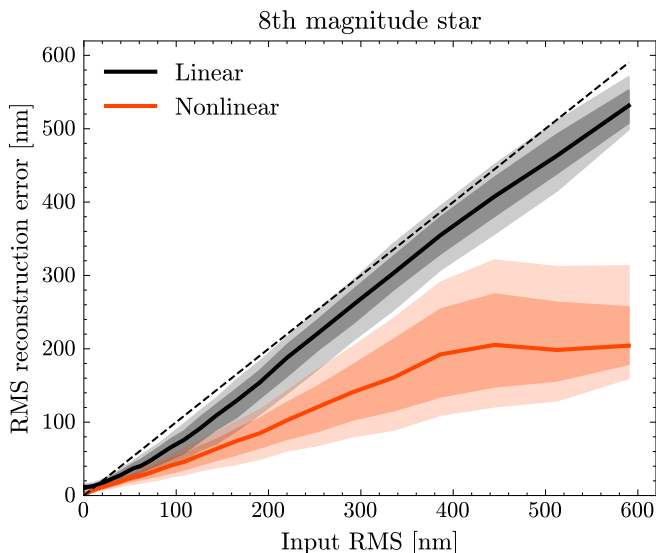
6

Figure 6.3: Comparison of the reconstruction accuracy for the nonlinear CNN with a linear model on the test dataset (same as Fig. 6.2), but now for an eighth magnitude star.

stellar magnitude to a total photon flux using the known zero point I-band magnitude for MagAO-X for the WFS arm [1].

We then used transfer learning to convert the model used in the previous section to one that can deal with these noisy measurements. This was done by training the model on the same dataset again, but now randomly sampling stellar magnitudes between 0 and 10. We then evaluated the open-loop reconstruction performance at the noise level for an eight magnitude star.

The optimal linear reconstructor also depends on the noise properties of the WFS measurements, and we therefore re-optimized the regularization strength for the linear model for this noise level. We found that 20 nm was within the noise limit, so we had to increase the maximum RMS that was used for the linear regression to 100 nm. The results are shown in Fig. 6.3, showing a decrease in performance of the nonlinear reconstructor on these noisy data. This is likely because it starts to struggle to distinguish between nonlinear structure and noise. This is in good agreement with our previous results (Landman & Haffert, 2020), in which we showed that the nonlinear correction gives the most improvement for bright stars. However, we still observe a major improvement over the linear model for large aberrations at this noise level.

---

[1]MagAO-x filter throughputs https://magao-x.org/docs/handbook/observers/filters.html

6.3.3 Sensitivity limit

To study the noise propagation properties of our reconstructor, we calculated the open-loop reconstruction performance for different stellar magnitudes. We then determined its sensitivity limit by finding the point at which the reconstruction RMS is 95% of the input RMS, that is, the limit of the smallest aberrations that can still be reconstructed. The resulting RMS sensitivity limit as a function of stellar magnitude is shown in Fig. 6.4. We show both the model trained on the noisy data ("Nonlinear faint") as well as the model only trained on the original high signal-to-noise ratio data ("Nonlinear bright"). The noise propagation of a Fourier-filtering WFS for a mode ϕ_i was derived in Chambouleyron et al. (2023). If dominated by photon noise, this is given by

$$\sigma^2(\phi_i) = \frac{1}{s_\gamma^2(\phi_i)N_{ph}}, \tag{6.4}$$

where $\sigma(\phi_i)$ is the RMS for mode ϕ_i, s_γ the photon noise sensitivity of the WFS to that mode and N_{ph} the total number of photons in the incoming beam. Assuming that the photon noise sensitivity for the unmodulated PWFS is roughly the same for each mode, the total RMS is given by

$$\sigma = \frac{1}{s_\gamma} \sqrt{\frac{N_{modes}}{N_{ph}}}, \tag{6.5}$$

where N_{modes} denotes the number of reconstructed modes. The unmodulated PWFS has $s_\gamma \approx 1.4$ (Guyon, 2005; Chambouleyron et al., 2023). The curve for this theoretical performance, assuming 1000 modes are reconstructed, is also shown in Fig. 6.4. Additionally, we recalibrated the linear model on the high signal-to-noise ratio data by only including input wavefronts up until 6 nm. This avoids the linear model being limited by nonlinearities and allows us to really obtain the sensitivity limit. Fig. 6.4 shows that for the brightest stars we are limited by systematics caused by bench turbulence, which is why the curve flattens of to a limit of ∼ 1.5 nm RMS. We see that both the linear model and the bright nonlinear model follow the same trend until a magnitude of ∼5, after which the CNN has much stronger noise amplification than the linear model. On the other hand, the nonlinear model trained on noisy data has better performance than the linear model and appears to even improve upon the analytical performance. We also see that it does not follow the classical $\propto 1/\sqrt{N_{ph}}$ curve. This is likely the result of the intrinsic regularization that is obtained by training the NN on noisy data. This effectively leads to the CNN being able to adapt the number of modes it reconstructs based on the WFS measurement. Furthermore, it can result in a biased estimator that, for example, always slightly underestimates noisy high-order

modes, thereby decreasing the noise propagation for these modes. The more advanced (nonlinear) regularization methods may help NNs deal with these noisy measurements, as was also noted in Wong et al. (2021). This added regularization comes at the cost of slightly decreased reconstruction accuracy for the brightest stars.
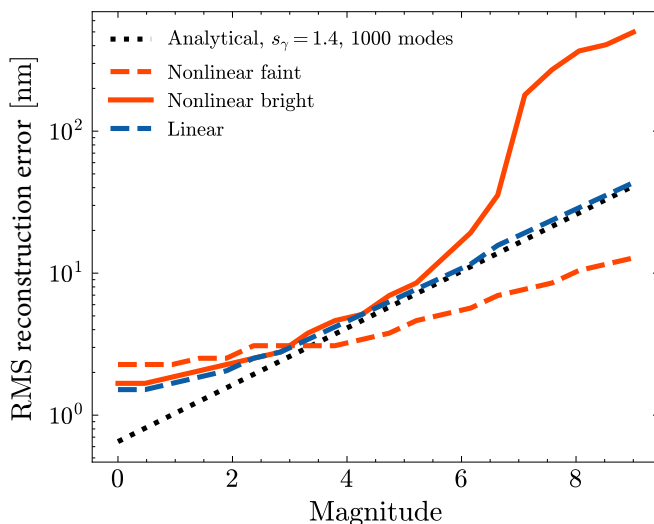

Figure 6.4: Smallest wavefront RMS that can be reconstructed as a function of the stellar I-band magnitude for the different reconstruction models. The analytical curve assumes a constant photon noise sensitivity of 1.4 for the unmodulated PWFS and the reconstruction of 1000 modes. The "Nonlinear bright" model refers to the model trained only on high signal-to-noise ratio data, while the "Nonlinear faint" model refers to the model that is also trained on noisy data.

6.4   Closed-loop tests

We evaluated the performance of the NN in closed loop for a variety of observing conditions. For all of these tests, we used a leaky integrator as our controller with a global gain (g) of 0.4 and leakage (l) of 1%, which are typical value used for MagAO-X on-sky, and the CNN only trained on the high signal-to-noise ratio data. For each iteration, the DM voltages were updated as follows:

$$\mathrm{DM}_{t+1} = (1 - l)\mathrm{DM}_t + gy_{\mathrm{pred}}. \tag{6.6}$$

Additionally, we used a separate tip-tilt loop as we observed small drifts in long-term tests. This was likely the result of tip-tilt drifts during the acquisition of the

training data and not explicitly using tip-tilt in our modal basis. We used a linear reconstructor for this tip-tilt loop with the total normalized intensity in each of the four pupils as inputs . This separate tip-tilt loop was run with a gain of 0.2 and leakage of 1%.

We generated atmospheric turbulence with statistics similar as expected at LCO using hcipy (Por et al., 2018) and the data from Prieto et al. (2010); Thomas-Osip et al. (2010); Males & Guyon (2018), and projected this on the modes spanned by the tweeter. This means higher order modes that cannot be produced by the tweeter are not present in these tests. Still, this includes a bit over 1000 more modes than the 1000 that we are controlling, indicating that our model is not influenced by the presence of higher-order modes in closed loop. We ran tests for seeing values of 0.55, 0.7 and 1.1 arcseconds, which correspond to slight better than median, slightly worse than median and 75% percentile conditions at LCO. As mentioned before, the tweeter cannot reproduce full atmospheric turbulence because of the limited number of modes. The fitting-error-limited Strehl using first-order estimates (Hardy, 1998) for these cases are 95%, 92%, and 85%, respectively. We neglected the effect of the actuator influence function during the projection. The shape of the DM influence functions reduces the amplitude of higher-order modes (Ruane et al., 2020), and this creates slightly better seeing conditions than expected. We also tested the impact of the effective wind speed on the closed-loop performance by testing two different values, the median wind speed with an C_n^2 weighted value of 18 ms^{-1} and double the median value at 36 ms^{-1}. While the tests were run using Python and not in real-time, we assumed an effective loop speed of 2 kHz in the simulation of the turbulence, which is the default loop speed for MagAO-X on bright stars (Males et al., 2022a). We repeated each of the experiments five times in order to test the stability and variance in performance.

The Strehl ratio was estimated from the focal plane images taken with the MagAO-X science camera. All Strehl measurements in this manuscript were done with the MagAO-X CH4 narrowband filter that has a center wavelength of 875 nm and a bandwidth of 26 nm [2]. We did this by calculating the encircled energy within a radius of 1 λ/D around the PSF peak. We normalized this by a reference Strehl measurement, which we obtained by closing the loop without applying any turbulence. The Strehl ratios shown throughout this work are relative to this reference measurement. We also found that the internal source had small drifts over time. To correct for this, we fitted a second order polynomial to the total intensity in the images over time and subsequently normalized the images. The resulting Strehl ratio curves are shown in Fig. 6.5. We observe an increase in

---

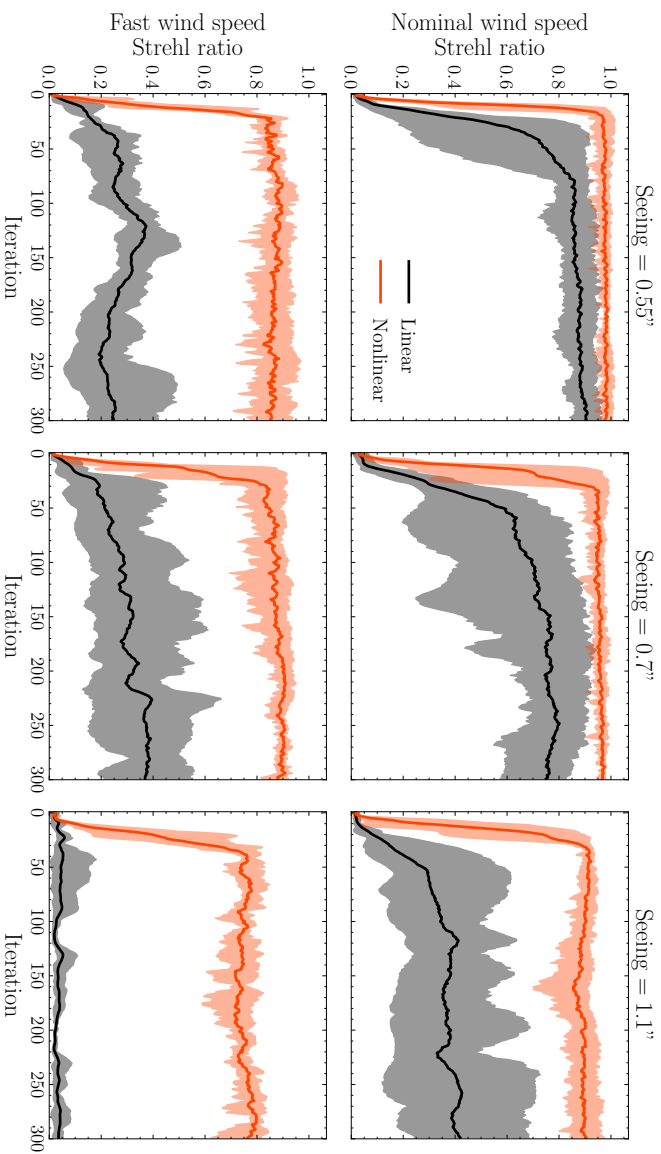[2]https://magao-x.org/docs/handbook/observers/filters.html

Figure 6.5: Closed-loop performance of the linear and nonlinear reconstructor for different atmospheric conditions. The solid line shows the mean Strehl at 875 nm over the five experiments, while the filled region indicates the minimum and maximum Strehl over the experiments. The top row shows the performance for a wind speed corresponding to median conditions at LCO with a C_n^2-weighted value of 18 ms^{-1}, while the bottom row uses double that value. The different columns correspond to different values for the seeing of the simulated turbulence. Videos of the closed-loop tests are available at https://zenodo.org/records/10580651.

Strehl ratio by using the nonlinear reconstructor over the linear reconstructor for each of the tested observing conditions. While the linear model is able to obtain $> 80\%$ only in good seeing conditions, the nonlinear model is able to do this even with 1.1" seeing. The Strehl that is achieved by the nonlinear model is in good agreement with the fitting-error-limited estimates. The difference between the reconstructors becomes even larger in the case of strong winds, as shown in the bottom row of Fig. 6.5. The nonlinearity of the PWFS leads to an underestimation of the wavefront with a linear model. If the change of the wavefront in a single iteration is larger than this correction, it is stuck in the highly nonlinear regime and is unable to converge to a high Strehl. On the other hand, the improved estimation of large aberrations with the nonlinear reconstructor leads to higher convergence rates and a better estimation of the residuals in closed loop.

6.4.1   Long-term stability

The long-term stability of nonlinear models in closed loop cannot be guaranteed. For example, modes to which the WFS is not sensitive can start to creep onto the DM, requiring a lower gain or higher leakage to ensure long-term stability. Furthermore, if the model is overly sensitive to the WFS camera or pupil alignment, it may not be stable over multiple days and might require recalibration often. We found this to not be a big issue in our case. For example, the tests shown in the previous section were conducted two days after the collection of the data that were used to train the model. This might not seem very long but after power cycling MagAO-X, its pupil always has to be realigned on the tweeter and a separate linear stage must be used to align the PWFS pupil on the WFS camera. These steps are currently done by hand, which means there is some amount of randomness in the alignment every time the system is powered on. Being able to close the loop with the CNN several days after the data were taken shows that it is quite robust against misalignment. We therefore do not expect to need many recalibrations of the CNN model. Even in the case of a required recalibration, we expect the recalibration to not take more than half an hour, as the previously calibrated model can be fine-tuned on newly collected data.

To test its long-term stability, we ran a closed-loop test for 2000 iterations. We observed stable behavior of the PSF and no modal creep. The total integrated PSF for this test, while excluding the first 100 iterations, is shown in Fig. 6.6. This shows the increased PSF stability with the nonlinear reconstructor as compared to the linear reconstructor, which will help boost exoplanet detection limits.
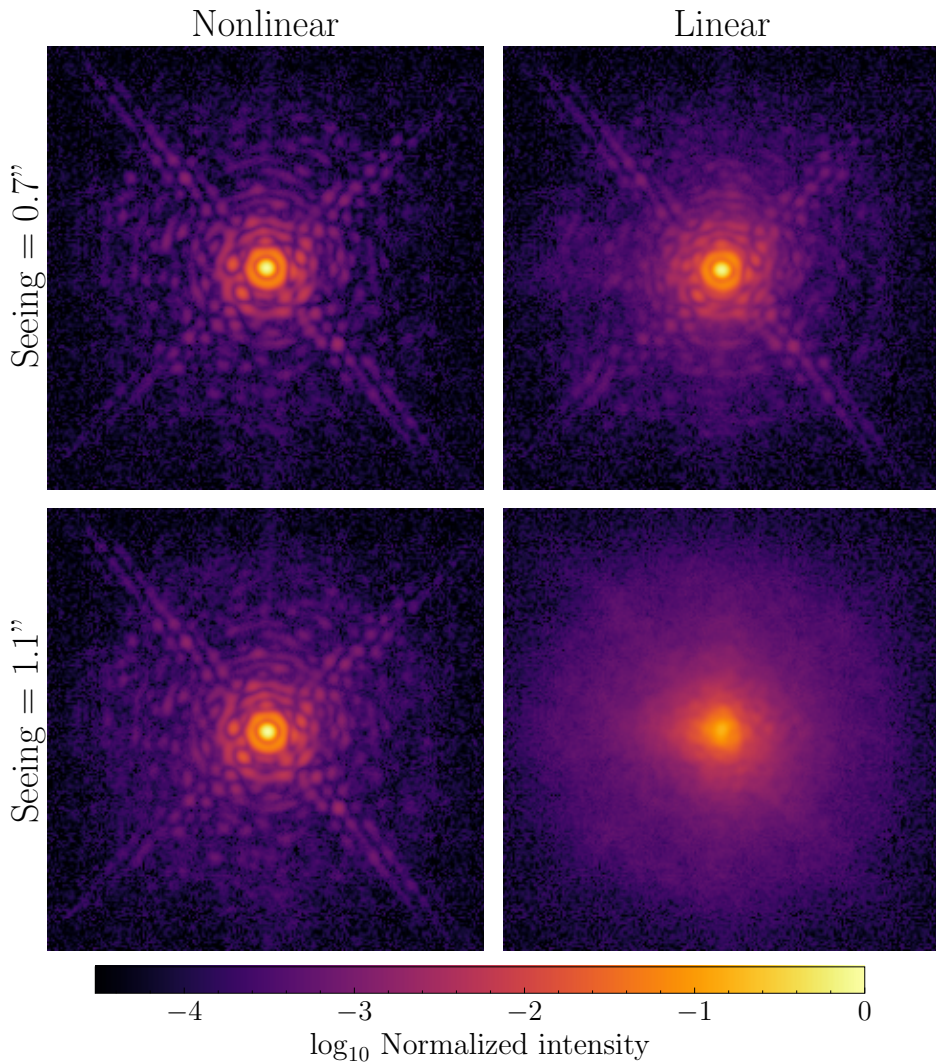
6

Figure 6.6: Integrated PSFs over a 2000-iteration closed-loop test for the nonlinear and linear reconstructor. The top row shows the performance for a seeing of 0.7", while the bottom row is for a seeing of 1.1". Videos of the closed-loop tests are available at `https://zenodo.org/records/10580651`.
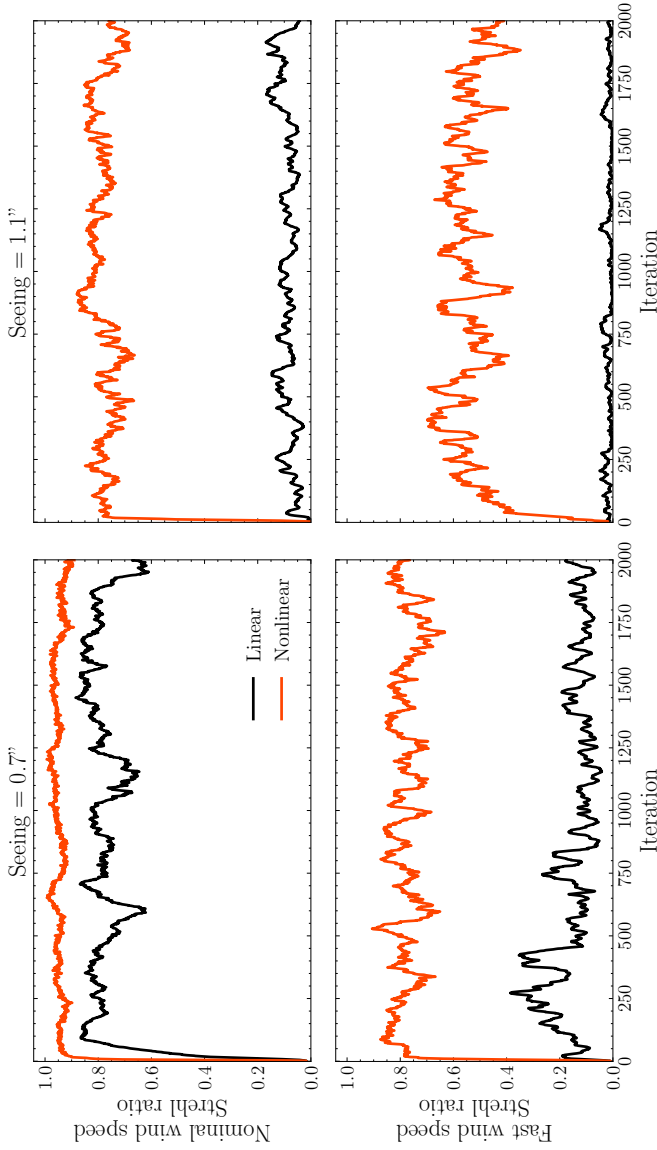
Figure 6.7: Closed-loop Strehl ratio for the linear and nonlinear reconstructors when the ND5 filter is used, decreasing the flux level by a factor of 10^3 (or a Δ mag = 7.5) compared to the data that were used to train the models. The different rows correspond to different values for the wind speed, while the columns show the performance for different seeing values. Videos of the closed-loop tests are available at https://zenodo.org/records/10580651.

6.4.2 Performance on fainter stars

Next, we tried to close the loop on a flux level that the CNN was not trained on. While Fig. 6.4 shows that the models can obtain good reconstruction accuracy over a large range of flux levels, the closed-loop stability is not necessarily guaranteed. We evaluated this by testing the "bright" model, which was only trained on high signal-to-noise ratio data, in closed loop with the ND5 filter. This reduces the flux level by a factor of 10^3 with respect to the training data and is equivalent to observing a ~7.5th magnitude star with MagAO-X, similar to the flux level for Proxima Centauri. We found that we did not have to tune the gain or leakage, as we observed stable behavior with the previously used values. The resulting Strehl ratios obtained during the closed-loop tests for different turbulence conditions is shown in Fig. 6.7. This shows a generally decreased performance of the AO system compared to the previous tests for both the linear and nonlinear model, which is due to the noisier WFS measurements. Still, the CNN has stable behavior in closed loop and can reach higher Strehl ratios than the linear reconstructor, showing that the nonlinear reconstructor can operate for a large range of stellar magnitudes. We note that Fig. 6.4 showed that the linear model has a better sensitivity limit for this stellar magnitude than the "bright" CNN, which we used for these test. However, the performance in closed-loop is still dominated by the nonlinearity error, as we are not operating around a diffraction-limited beam. This explains the improved performance of the CNN over the linear model, even though it has stronger noise propagation.

6.4.3 Inference time

Finally, we tested the inference speed of our model. We converted the trained model to a TensorRT optimized model with half precision, and tested the models on a single GeForce RTX 2080 Ti GPU. To remove the effect of overheads such as data transfer to the GPU on these measurements, we compared the inference speed of our models to an empty model, which returns a constant and does not do any processing. We find that reconstruction of a single WFS measurement takes on average $690 \pm 50 \ \mu s$, compared to $170 \pm 50 \ \mu s$ for the linear model. This means that the inference is too slow to be run at 2 kHz, but loop speeds of >1 kHz are feasible. Pruning of the model could be used to further increase the inference speed (Asif et al., 2019). Alternatively, the size of the nonlinear model could be decreased. This comes at a slight decrease in reconstruction accuracy in the nonlinear regime, but in our experience this decrease is relatively small. Integration of TensorRT within the MagAO-X software environment, which is based on CACAO (Guyon et al., 2018), is currently ongoing. A true test of the

inference speed and jitter of the model will be conducted after this integration and the model will be adjusted such that the standard 2 kHz loop speed can be achieved.

6.5 Conclusions

We have presented the first closed-loop lab demonstration of a CNN-based reconstructor for the unmodulated PWFS with MagAO-X. Our nonlinear reconstructor has a significantly improved dynamic range of >600 nm RMS compared to the ~ 50 nm RMS dynamic range for classical linear reconstruction. While its ability to accurately reconstruct the wavefront in the nonlinear regime decreases for fainter stars, we still observed a major improvement over a linear reconstructor. We have shown that the nonlinear reconstructor can reach the sensitivity limit of the PWFS and does not lead to stronger noise amplification when including noisy data in the training process. In this case, the nonlinear model does not obey the standard noise propagation scaling due to its intrinsic regularization properties. Closed-loop tests confirmed the increased dynamic range, with the nonlinear reconstructor reaching higher Strehl ratios than a classical linear reconstructor. The improved performance is the most pronounced when conditions are suboptimal, as the linear reconstructor is not able to converge in closed loop for the worst simulated conditions. We have extensively tested the stability of the nonlinear model and found that it is stable over multiple days, in long-term closed-loop tests and when tested on a different flux level than the one it was originally trained on. The presented work demonstrates that it is possible to use the unmodulated PWFS in most atmospheric conditions and that it might not be necessary to modulate the PWFS anymore in the future.

The obvious next step is to test the nonlinear reconstructor on-sky. Current work focuses on decreasing the computational complexity to reach 2 kHz speeds and implementing the TensorRT framework within the MagAO-X software environment to enable on-sky testing. This will make MagAO-X a pathfinder AO system for testing nonlinear control with (deep) NNs.

6

References

Agapito, G., Pinna, E., Esposito, S., Heritier, C. T., & Oberti, S. 2023, A&A, 677, A168, doi: `10.1051/0004-6361/202346359`

Archinuk, F., Hafeez, R., Fabbro, S., Teimoorinia, H., & Véran, J.-P. 2023, Journal of Astronomical Telescopes, Instruments, and Systems, 9, 049005, doi: `10.1117/1.JATIS.9.4.049005`

Asif, U., Tang, J., & Harrer, S. 2019, arXiv e-prints, arXiv:1909.08097, doi: `10.48550/arXiv.1909.08097`

Bertrou-Cantou, A., Gendron, E., Rousset, G., et al. 2022, A&A, 658, A49, doi: `10.1051/0004-6361/202141632`

Boccaletti, A., Chauvin, G., Wildi, F., et al. 2022, in Ground-based and Airborne Instrumentation for Astronomy IX, Vol. 12184, SPIE, 622–634

Bond, C. Z., Sauvage, J.-F., Schwartz, N., et al. 2022, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12185, Adaptive Optics Systems VIII, ed. L. Schreiber, D. Schmidt, & E. Vernet, 121851B, doi: `10.1117/12.2627713`

Bowler, B. P. 2016, PASP, 128, 102001, doi: `10.1088/1538-3873/128/968/102001`

Chambouleyron, V., Fauvarque, O., Janin-Potiron, P., et al. 2020, A&A, 644, A6, doi: `10.1051/0004-6361/202037836`

Chambouleyron, V., Fauvarque, O., Plantet, C., et al. 2023, A&A, 670, A153, doi: `10.1051/0004-6361/202245351`

Chambouleyron, V., Fauvarque, O., Sauvage, J. F., Neichel, B., & Fusco, T. 2021, A&A, 649, A70, doi: `10.1051/0004-6361/202140354`

Chambouleyron, V., Sengupta, A., Salama, M., et al. 2024, A&A, 681, A48, doi: `10.1051/0004-6361/202347220`

Close, L. M., Males, J. R., Durney, O., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10703, Adaptive Optics Systems VI, ed. L. M. Close, L. Schreiber, & D. Schmidt, 107034Y, doi: `10.1117/12.2312280`

Deo, V., Gendron, É., Rousset, G., et al. 2019, A&A, 629, A107, doi: `10.1051/0004-6361/201935847`

Deo, V., Gendron, É., Vidal, F., et al. 2021, A&A, 650, A41, doi: `10.1051/0004-6361/202040216`

Engler, B., Le Louarn, M., Vérinaud, C., Weddell, S., & Clare, R. 2022, Journal of Astronomical Telescopes, Instruments, and Systems, 8, 021502, doi: `10.1117/1.JATIS.8.2.021502`

Esposito, S., & Riccardi, A. 2001, A&A, 369, L9, doi: `10.1051/0004-6361:20010219`

Fitzsimmons, J., Agapito, G., Bonaglia, M., et al. 2020, in Adaptive Optics Systems VII, Vol. 11448, SPIE, 1299–1311

Frazin, R. A. 2018, Journal of the Optical Society of America A, 35, 594, doi: `10.1364/JOSAA.35.000594`

Guyon, O. 2005, The Astrophysical Journal, 629, 592

Guyon, O. 2018, ARA&A, 56, 315, doi: `10.1146/annurev-astro-081817-052000`

Guyon, O., Sevin, A., Gratadour, D., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10703, Adaptive Optics Systems VI, ed. L. M. Close, L. Schreiber, & D. Schmidt, 107031E, doi: `10.1117/12.2314315`

6

Haffert, S. Y., Males, J. R., Close, L. M., et al. 2022, in Adaptive Optics Systems VIII, Vol. 12185, SPIE, 1071–1079

Haffert, S. Y., Males, J. R., Close, L. M., et al. 2021, Journal of Astronomical Telescopes, Instruments, and Systems, 7, 029001, doi: `10.1117/1.JATIS.7.2.029001`

Hardy, J. W. 1998, Adaptive optics for astronomical telescopes, Vol. 16 (Oxford Optical and Imaging Sci)

Hedglen, A. D., Close, L. M., Haffert, S. Y., et al. 2022, Journal of Astronomical Telescopes, Instruments, and Systems, 8, 021515, doi: `10.1117/1.JATIS.8.2.021515`

Hutterer, V., Neubauer, A., & Shatokhina, J. 2023, Inverse Problems, 39, 035007, doi: `10.1088/1361-6420/acb568`

Hutterer, V., & Ramlau, R. 2018, Appl. Opt., 57, 8790, doi: `10.1364/AO.57.008790`

Kasper, M., Cerpa Urra, N., Pathak, P., et al. 2021, The Messenger, 182, 38, doi: `10.18727/0722-6691/5221`

Kingma, D. P., & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980. `https://arxiv.org/abs/1412.6980`

Landman, R., & Haffert, S. Y. 2020, Optics Express, 28, 16644, doi: `10.1364/OE.389465`

Landman, R., Haffert, S. Y., Radhakrishnan, V. M., & Keller, C. U. 2021, Journal of Astronomical Telescopes, Instruments, and Systems, 7, 039002, doi: `10.1117/1.JATIS.7.3.039002`

Landman, R., Keller, C., Por, E. H., et al. 2022, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12185, Adaptive Optics Systems VIII, ed. L. Schreiber, D. Schmidt, & E. Vernet, 1218589, doi: `10.1117/12.2629930`

Males, J. R., & Guyon, O. 2018, Journal of Astronomical Telescopes, Instruments, and Systems, 4, 019001

Males, J. R., Close, L. M., Miller, K., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10703, Adaptive Optics Systems VI, ed. L. M. Close, L. Schreiber, & D. Schmidt, 1070309, doi: `10.1117/12.2312992`

Males, J. R., Close, L. M., Haffert, S., et al. 2022a, in Adaptive Optics Systems VIII, Vol. 12185, SPIE, 61–70

Males, J. R., Close, L. M., Haffert, S. Y., et al. 2022b, in Adaptive Optics Systems VIII, Vol. 12185, SPIE, 1381–1391

Nousiainen, J., Rajani, C., Kasper, M., et al. 2022, A&A, 664, A71, doi: `10.1051/0004-6361/202243311`

Perera, S., Maire, J., Do Ó, C. R., et al. 2022, in Adaptive Optics Systems VIII, Vol. 12185, SPIE, 1298–1304

Pinna, E., Esposito, S., Hinz, P., et al. 2016, in Adaptive Optics Systems V, Vol. 9909, SPIE, 1183–1192

Por, E. H., Haffert, S. Y., Radhakrishnan, V. M., et al. 2018, in Proc. SPIE, Vol. 10703, Adaptive Optics Systems VI, doi: `10.1117/12.2314407`

Pou, B., Smith, J., Quinones, E., Martin, M., & Gratadour, D. 2022, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12185, Adaptive Optics Systems VIII, ed. L. Schreiber, D. Schmidt, & E. Vernet, 121852U, doi: `10.1117/12.2627849`

Prieto, G., Thomas-Osip, J. E., Phillips, M. M., McCarthy, P., & Johns, M. 2010, in

6

Ground-based and Airborne Telescopes III, Vol. 7733, SPIE, 1616–1625

Ragazzoni, R. 1996, Journal of Modern Optics, 43, 289, doi: `10.1080/09500349608232742`

Ragazzoni, R., & Farinato, J. 1999, A&A, 350, L23

Ronneberger, O., Fischer, P., & Brox, T. 2015, arXiv e-prints, arXiv:1505.04597. `https://arxiv.org/abs/1505.04597`

Ruane, G., Echeverri, D., Bendek, E., et al. 2020, Journal of Astronomical Telescopes, Instruments, and Systems, 6, 045002

Schatz, L. H., Males, J. R., Close, L. M., et al. 2018, in Adaptive Optics Systems VI, Vol. 10703, SPIE, 671–679

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, Journal of Machine Learning Research, 15, 1929. `http://jmlr.org/papers/v15/srivastava14a.html`

Swanson, R., Lamb, M., Correia, C., Sivanandam, S., & Kutulakos, K. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10703, Adaptive Optics Systems VI, ed. L. M. Close, L. Schreiber, & D. Schmidt, 107031F, doi: `10.1117/12.2312590`

Thomas-Osip, J. E., Prieto, G., McWilliam, A., et al. 2010, in Ground-based and Airborne Telescopes III, Vol. 7733, SPIE, 1606–1615

Wong, A. P., Norris, B. R. M., Tuthill, P. G., et al. 2021, Journal of Astronomical Telescopes, Instruments, and Systems, 7, 019001, doi: `10.1117/1.JATIS.7.1.019001`

Wong, A. P., Norris, B. R. M., Deo, V., et al. 2023, PASP, 135, 114501, doi: `10.1088/1538-3873/acfdcb`

6