Tempestas ex machina: A review of machine learning methods for wavefront control

J. Fowler[a] and Rico Landman[b]

[a]Department of Astronomy & Astrophysics, University of California, Santa Cruz, CA, USA
[b]Leiden Observatory, Leiden University, Leiden, The Netherlands

ABSTRACT

As we look to the next generation of adaptive optics systems, now is the time to develop and explore the technologies that will allow us to image rocky Earth-like planets; wavefront control algorithms are not only a crucial component of these systems, but can benefit our adaptive optics systems without requiring increased detector speed and sensitivity or more effective and efficient deformable mirrors. To date, most observatories run the workhorse of their wavefront control as a classic integral controller, which estimates a correction from wavefront sensor residuals, and attempts to apply that correction as fast as possible in closed-loop. An integrator of this nature fails to address temporal lag errors that evolve over scales faster than the correction time, as well as vibrations or dynamic errors within the system that are not encapsulated in the wavefront sensor residuals; these errors impact high contrast imaging systems with complex coronagraphs. With the rise in popularity of machine learning, many are investigating applying modern machine learning methods to wavefront control. Furthermore, many linear implementations of machine learning methods (under varying aliases) have been in development for wavefront control for the last 30-odd years. With this work we define machine learning in its simplest terms, explore the most common machine learning methods applied in the context of this problem, and present a review of the literature concerning novel machine learning approaches to wavefront control.

Keywords: machine learning, wavefront control, adaptive optics, predictive control, linear quadratic Gaussian control (LQG), neural networks, reinforcement learning, Kalman filtering, empirical orthogonal functions (EOF)

1. INTRODUCTION

In the search for and characterization of extrasolar planets, direct imaging is a powerful tool to unlock the detection of planets not available via other detection methods (i.e., planets that will not geometrically transit) and characterize them without having to disentangle the light of the planet from its host star. From the ground, direct imaging is only possible with extreme adaptive optics (exAO) where a very high level of correction over a small field of view enables high contrast imaging. Only stellar point spread functions (PSF) with extremely high Strehl ratios will enable coronagraphy to effectively null host starlight, and result in high enough contrasts to resolve the light of substellar companions. The scope of this review is limited specifically to wavefront control methods used within the field of ground-based adaptive optics for astronomical instrumentation, with a keen interest in this extreme adaptive optics context. (For a review of machine learning methods for wavefront sensing, see Wong, 2021.[1])

An adaptive optics system at its simplest is made of 3 fundamental components: a wavefront sensor (which records information on the state of the wavefront after it has been aberrated by Earth-atmosphere), a deformable mirror (which attempts to correct for those aberrations), and a control algorithm that determines what correction the deformable mirror (DM) should apply given the wavefront sensor information available. Figure 1 shows a control diagram of this system.

To date, every extreme AO system runs their standard control for mid and high order modes as an integrator. This refers specifically to a closed-loop integral style controller, where the wavefront sensor is downstream of the

---

Further author information: (Send correspondence to J.F.)
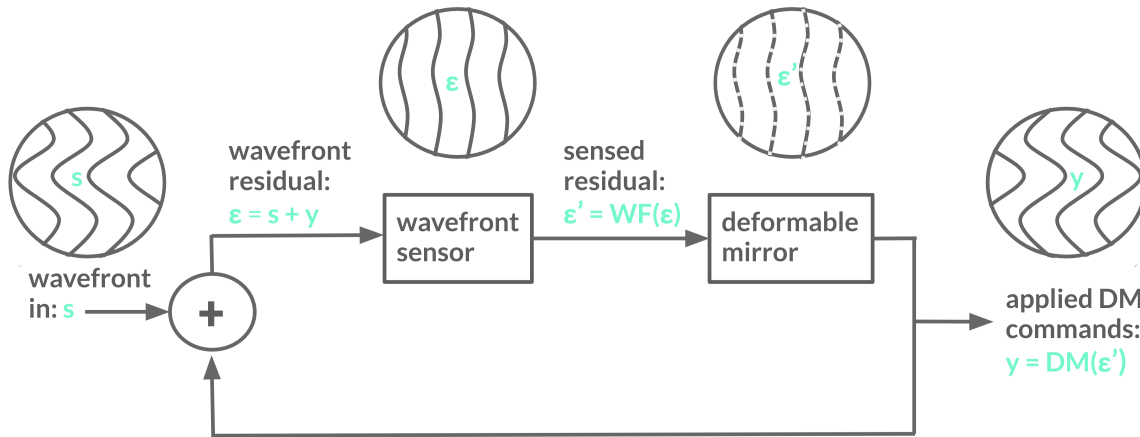J.F.: E-mail: jumfowle@ucsc.edu

Figure 1. The full state of turbulence **s** enters the system and immediately encounters the wavefront correction **y**, noted by the sum junction. The wavefront sensor senses the residual wavefront $\boldsymbol{\epsilon} = \mathbf{s} + \mathbf{y}$, and reports a digitized and imperfect realization $\boldsymbol{\epsilon}'$. Finally, a computer uses this information to calculate the command update to the deformable mirror, which feeds the next iteration.

DM, senses residual wavefront after the correction, estimates the error with a simple minimization between that residual and a flat reference, and applies that command as an addition to the current shape with some gain that is typically between $0.2 < g < 0.6$, or a gain optimized per mode.[2,3] This correction is applied as quickly as possible, with the hope that the speed of the system will keep up with the speed of the evolving atmosphere. In practice this lag-time ranges from 1-5 ms.

However, not only do we want to optimize our current observations with large telescopes like VLT, Subaru, Magellan, and Keck, as we look forward to ELTs, we will require higher fidelity correction to detect and characterize earth-like planets around sun-like stars. For example, detecting oxygen in the atmosphere of Proxima Centauri b will require a raw contrast of 3×10^{-5} in the visible at $10~\lambda/\mathrm{D}$,[4] something our current control algorithms and systems have yet to accomplish. Much of this is limited by technology, but some of it is limited by the performance of the control algorithm. Furthermore, as we scale up to larger systems, with higher actuator count deformable mirrors and wavefront sensors with many more subapertures, the act of calibrating the DM to the wavefront sensor poke by poke or pixel by pixel becomes considerably more time-consuming, and a problem that is interesting to consider offloading to a machine learning algorithm.

We expect the predictable elements of atmospheric physics are governed entirely by linear systems; state of the art atmospheric modelling for adaptive optics is done with Kolmogorov turbulence[5,6] that evolves according to the frozen flow hypothesis,[7] where velocities imparted by wind layers are expected to be the driving factor. This model has been confirmed from the perspective of an AO system.[8] Additional atmospheric turbulence (i.e., boiling,[9] dome seeing[10]) adds random features that could still be statically encapsulated with linear noise propagation (i.e., Kalman filtering and AR models.)

Wavefront sensing is an interesting potential source of non-linearities, covered at length in Wong, 2022.[1] In short, wavefront sensing often operates as a tradeoff between sensitivity and linear sensible range (the more sensitive the wavefront sensor is, the lower the usable linear range becomes.) While we attempt to design our wavefront sensor in such a way that this assumption of linearity is valid (i.e. by introducing dynamic modulation), the Pyramid and Zernike wavefront sensors are known to have nonlinear responses for large aberrations. However, running wavefront control in closed-loop leaves residual turbulence sensed by the wavefront sensor within the linear range of most wavefront sensors used for first-stage AO. Similarly, deformable mirrors (DMs) operate with actuator pokes that have a limited linear range; in practice, the linear range of the DM is much larger than that of the wavefront sensor.

In order to comment on and evaluation the performance of control methods, we first set priorities for current and future instruments. We define mandatory criteria and ideal exploration space for new methods within an extreme exAO system in Table 1. As we summarize these methods, we examine which methods achieve some of these ideal features and whether or not they sacrifice mandatory AO requirements.

Table 1. Here we define criteria deemed mandatory for extreme AO systems (i.e., any method in development cannot sacrifice these), and criteria that we could potentially gain with future technology developments.

mandatory criteria	
	achieves at least 60% Strehl ratio in the IR
	runs in real time
	robust over a full night
	accounts for static NCPAs
ideal criteria	
	accounts for non-linearities
	does not require high fidelity instrument model
	high sky coverage at high performance
	tunable and flexible
	accounts for *dynamic* NCPAs
	does not require consistent operator attention

2. MACHINE LEARNING METHODS COMMON TO THE FIELD

2.1 Formal Definitions of Machine Learning

Despite recent explosions in complex machine learning and artificial intelligence algorithms meant to write software, have a conversation, or win chess games, for this review a more basic definition of machine learning will suffice. We define machine learning as any problem that uses one set of data to make rules for a second set of data.[11] From this definition, machine learning not only includes complicated many-layer neural networks and artificial intelligence engines, but also classic linear regression problems, extrapolation problems, and many basic control engineering and statistical concepts (i.e., Kalman filtering and linear quadratic Gaussian controllers.) In this section, we will overview the most common machine learning methods applied to wavefront control, as well as comment on the overlap in methods.

2.2 Kalman Filtering and Linear Quadratic Gaussian Control

2.2.1 Kalman filtering

Kalman filtering[12] is a statistical prediction tool, common to engineering and mathematics alike. Simply stated, a Kalman filter is a robust way to predict the evolution of a property according to a physical law, a Gaussian uncertainty distribution, and data fed back into the system to verify and improve that estimate. Figure 2 shows a visual representation of how some property (for our purposes, the measurement from a wavefront sensor) evolves according to some physical law (for our purposes, the wind velocity affecting the atmospheric turbulence, any dynamics in the system, and the control law we are applying), and a cloud of uncertainty that accompanies and is altered by that prediction.

A Kalman filter starts with a state vector $\hat{\mathbf{x}}$, which encapsulates all the variables that we predict at each iteration, and a covariance matrix $\mathbf{P}$, which uses the covariance to estimate the uncertainty of those measurements. Given a physical law $\mathbf{F}$, a control law $\mathbf{B}$, and a noise term $\mathbf{Q}$, we predict the next state vector and covariance matrix with
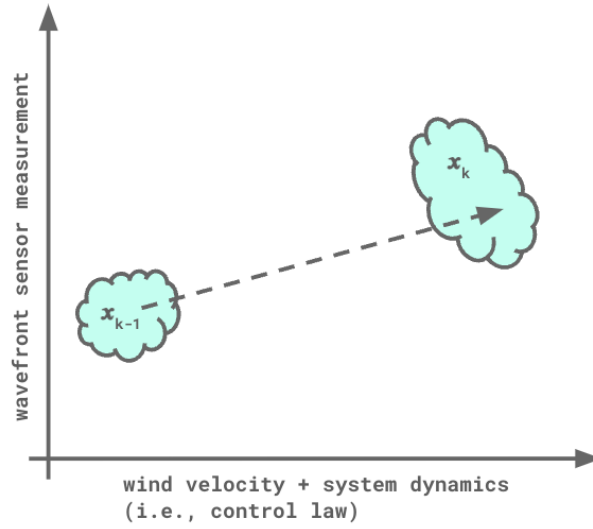
Figure 2. "A Kalman filter is used to generate estimates of the states, given the noise-corrupted measurements."[13] A Kalman filter projects a state vector and a Gaussian uncertainty cloud forward according to a physical law.

$$\hat{\mathbf{x}}_\mathbf{k} = \mathbf{F_k}\hat{\mathbf{x}}_{\mathbf{k-1}} + \mathbf{B_k}\vec{\mathbf{u}_k} \qquad (1)$$

$$\mathbf{P_k} = \mathbf{F_k}\mathbf{P_{k-1}}\mathbf{F_k^T} + \mathbf{Q_k} \qquad (2)$$

Despite a simple step-by-step update formula, a Kalman filter is not typically updated at each iteration in practice. Instead, a steady-state solution is found after training from previous system data.

2.2.2 Linear quadratic Gaussian control

A common pair in control engineering is a Kalman filter and a linear quadratic regulator, or a controller that aims to minimize a quadratic cost function at each iteration. This combination is referred to as a linear quadratic Gaussian (LQG) controller; a controller that assumes Gaussian errors, predicts forward with a linear physical law, and regulates the system with a quadratic cost function. Figure 3 shows an example of an LQG controller in the context of wavefront control.

$$\mathbf{J} = \mathbf{E}\left\langle \mathbf{\Sigma}_{i=0}^{N}\frac{1}{2}\left[\mathbf{x^T}(t_i)\mathcal{X}\mathbf{x}(t_i) + \mathbf{u^T}(t_i)\mathcal{U}\mathbf{u}(t_i)\right]\right\rangle \qquad (3)$$

$\mathbf{J}$ is an example cost function[13] used for a linear quadratic regulator, given some state vector $\mathbf{x}$, deformable mirror commands $\mathbf{u}$, and expectation value $\mathbf{E}$. $\mathbf{J}$ minimizes the error in the state as well as DM commands to estimate the next DM command.

2.3 Empirical Orthogonal Functions

Empirical orthogonal functions (EOF)[14] estimates the future of a wavefront by using past data to find linear temporal trends. It is basis independent (it can be used with zonal modes, Fourier models, KL modes, or any mix that allows for robust non-competing control), though in practice it has been mostly implemented with zonal (one data point per wavefront sensor subaperture/DM actuator) modes.

First, we build a history vector $\mathbf{h}$(t), that contains some number of associated wavefront sensor states. Depending on the implementation, this vector can be per mode (i.e. for m modes in a basis, $w_0, w_1, ...w_{m-1}$), or $\mathbf{h}_m$, mapping to a specific mode. Equation 4 shows the example for all modes and many states in time.
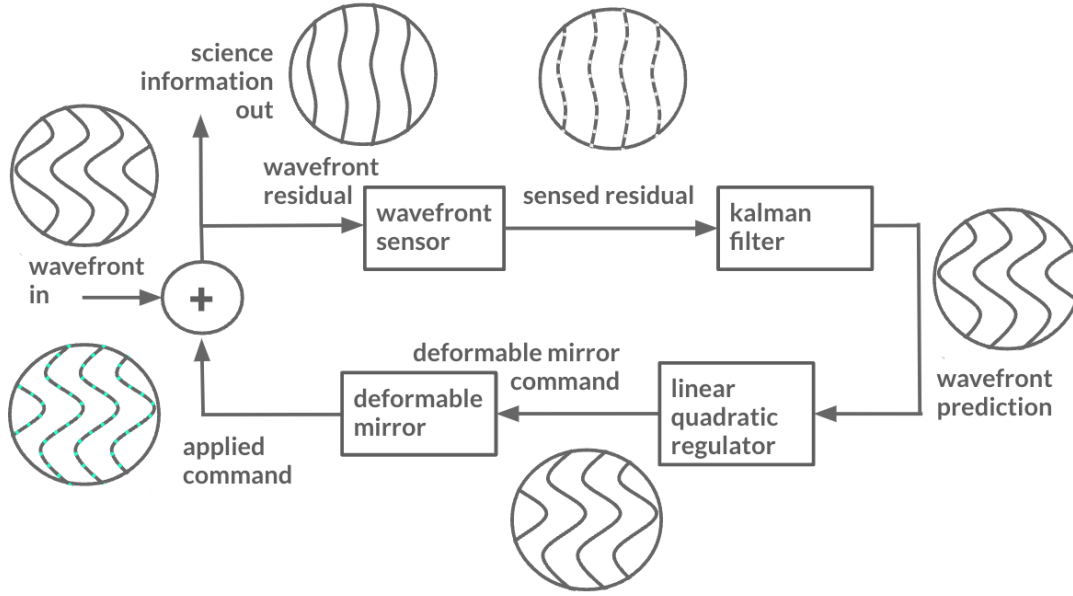
Figure 3. Control diagram of a linear quadratic Gaussian (LQG) controller. In this system, a sensed wavefront residual is fed to a Kalman filter, which predicts the future state of the wavefront, calculates the control law with a linear quadratic regulator, and sends that control command to the deformable mirror to correct the system.

$$
\mathbf{h}(t) = \begin{bmatrix} \mathbf{w}_0(t) \\ \mathbf{w}_1(t) \\ ... \\ \mathbf{w}_{m-1}(t) \\ \mathbf{w}_0(t - dt) \\ ... \\ \mathbf{w}_{m-1}(t - (n-1)dt) \end{bmatrix}
\tag{4}
$$

The goal of EOF is to build a predictive filter $\mathbf{F}$ (or $\mathbf{F}_m$ if it maps to a single mode) that will predict a future wavefront, such that some **predicted wavefront** = $\mathbf{Fh}$. This predictive filter can be calculated every iteration, or updated every time the linear evolution of the turbulence changes (which we would roughly expect to follow evolving wind velocity over the course of a night.)

$$
\mathbf{Fh}(t) = \begin{bmatrix} \mathbf{w}_0(t + dt) \\ ... \\ \mathbf{w}_{m-1}(t + dt) \end{bmatrix}
\tag{5}
$$

To find the predictive filter, we build $\mathbf{D}$, a matrix which holds many history vectors that act as training data. They will be minimized against a truth condition $\mathbf{P}$, which holds the correct prediction corresponding to each history vector. This leads to a simple minimization problem

$$
\min_{\mathbf{F}} ||\mathbf{D}^T\mathbf{F}^T - \mathbf{P}^T||^2.
\tag{6}
$$

The solution for the predictive filter can be found using a simple pseudo-inverse:

$$
\mathbf{F} = ((\mathbf{D}^T)^\dagger \mathbf{P}^T)^T.
\tag{7}
$$

This inversion can be done in a handful of ways; Guyon 2017[14] used a singular value decomposition (SVD), but other works[15–17] have used a least squares inversion.

EOF as a predictor works so simply because we consider some wavefront $\mathbf{w}_m(\mathrm{t})$ as the full state of turbulence that the atmosphere imparts. In practice, this is not the measurement we have in a typical AO system, which runs in closed-loop, where our wavefront sensor will measure a residual after some DM correction has been applied. In this implementation, EOF uses pseudo-open loop (POL) reconstructed values of the full wavefront, where information on the applied DM commands and measured wavefront sensors residuals are used to reconstruct what the full state of the turbulence would have been. This is how this is run on-sky both at Subaru/SCExAO[18] and Keck/NIRC2.[16]

However, the math represented by EOF is arguably comparable to that of Dessenne, 1998,[19] which outlined a similar method, but builds a modified history vector using DM commands and wavefront sensor slopes, instead of reconstructing the full state of the turbulence for each input mode. This implementation is more computationally expensive, but runs in closed-loop.

2.4 Neural Networks

The previous methods all rely on a linear state space model; the predicted future state can be expressed as a linear combination of previous states and commands. The advantage of this linear model is that one can often find analytical solutions that minimize a quadratic cost function, as discussed in the previous section. However, this assumption of linearity is not valid in all cases. An alternative is to use a nonlinear state-space model. This can be done using Neural Networks (NN), which can act as arbitrary function approximators. These NN's utilize multiple consecutive matrix multiplications with a so-called activation function in between. This activation function is nonlinear, allowing the NN to model nonlinear functions. These NN's are often visualized by having multiple "hidden" layers between the input and output, where the activation $\mathbf{y}$ in layer i can be mathematically expressed as:

$$\mathbf{y}_i = f(W\mathbf{y}_{i-1} + \mathbf{b}). \tag{8}$$

Here, W are the weights of that layer, $\mathbf{b}$ the bias, and $f()$ the activation function. A visualization of a NN is shown in Fig. 5. Note that a single-layer NN without an activation function reduces to a familiar linear model, such as in EOF. The amount of hidden layers and number of neurons in each layer is tweaked based on the complexity of the problem. The weights and biases can be optimized by minimizing a chosen loss function on a training dataset. This is commonly done using stochastic gradient descent based optimizers in combination with the backpropagation algorithm. Libraries such as `keras` and `pytorch` hide most of these complexities to end users, which makes building and training basic NN's much easier.

2.4.1 Advanced architectures

For complex nonlinear problems basic NN's quickly run into the curse of dimensionality, where the number of free parameters explodes. Luckily, the parameterization of these NN's can be adapted to specific problems by using domain knowledge. This can be seen as a strong form of regularization on the complexity of the model. For example, Convolutional Neural Networks (CNN's) are popular for image processing tasks. They use the knowledge that features are local and translational invariant to drastically reduce the number of free parameters that have to be optimized. For our purposes, CNN's can be used to capture spatial correlation in the wavefront. Alternatively, Recurrent Neural Networks (RNN's) are commonly used for problems involving sequences. They are well suited to deal with the temporal aspect of predictive control. Popular variants of RNN's are Long-short term memory (LSTM) networks and Gated Recurrent Units (GRU).

2.5 Reinforcement Learning

Reinforcement Learning (RL) is a fast growing field that deals with learning optimal control through interaction with environment. The goal of these algorithms is to find the optimal commands that maximize some arbitrary optimization objective. RL algorithms can be split up into two main branches: model-free and model-based. Note that models are used in both cases.
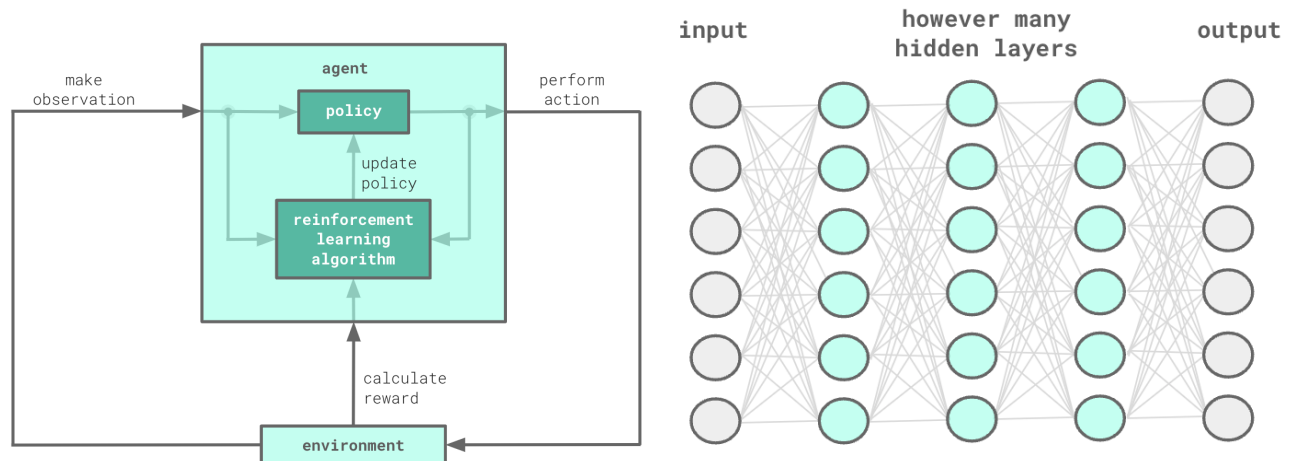
Figure 4.
Left: Diagram of a reinforcement learning (RL) algorithm. An RL algorithm enacts some policy (control law) on some environment, makes observations, optimizes itself according to a reward. Right: Diagram of a neural network (NN). A neural network finds non-linear connections between layers of nodes.

2.5.1 Model-based reinforcement learning

Model-based RL algorithms explicitly learn a state-space or dynamics model. Using this state-space model and the currently measured state, they aim to find the control command which maximizes the optimization objective at each iteration. This can be done using planning algorithms, which can be quite computionally expensive to run at kHz frequencies. In the case of a linear dynamics model with a quadratic cost function, model-based RL reduces to the LQG problem with online identification of the state-space model.

2.5.2 Model-free Reinforcement Learning

Instead of learning a dynamics model, model-free RL algorithms aim to directly learn an optimal parameterized control law. Popular choices of model-free RL are variants of the policy gradient algorithm, such as Soft Actor-Critic (SAC) and Deep Deterministic Policy Gradient (DDPG). All these algorithms try to learn a differentiable model of the optimization metric as a function of the state and control command. One can then obtain the gradient of the cost function with respect to the control commands. This can subsequently be used to optimize the parameters of the control law. Model-free RL methods provide a parameterized control law, with which inference speeds are typically fast enough to run at kHz frequencies. A disadvantage of model-free algorithms are that they are typically less data efficient and would thus more slowly adapt to changing conditions. Promising hybrid methods, that combine the advantages of model-free and model-based algorithms, are also being developed.

2.6 Overlaps Between Fields

In summary, four major concepts will be discussed in the rest of this review. Reinforcement learning and neural networks are two common elements of machine learning that often go hand-in-hand: reinforcement learning as a particular problem type, and neural networks as a common solver for that problem type.

Narrowing the scope of these problems and making them linear shows that an LQG controller is simply a reinforcement learning problem with more conditions about its error propagation, physical law, and cost function. In particular, a mathematical analysis of the overlap between a reinforcement learning problem and an LQG controller is available in Recht, 2018.[20] Kalman filtering is often utilized as an element of an LQG controller, though one of the methods we describe is closer to a pure Kalman filter approach. Finally, EOF acts as a one-layer neural network, and works that have compared EOF to neural networks[21,22] find that they converge to the EOF solution for linear problems.
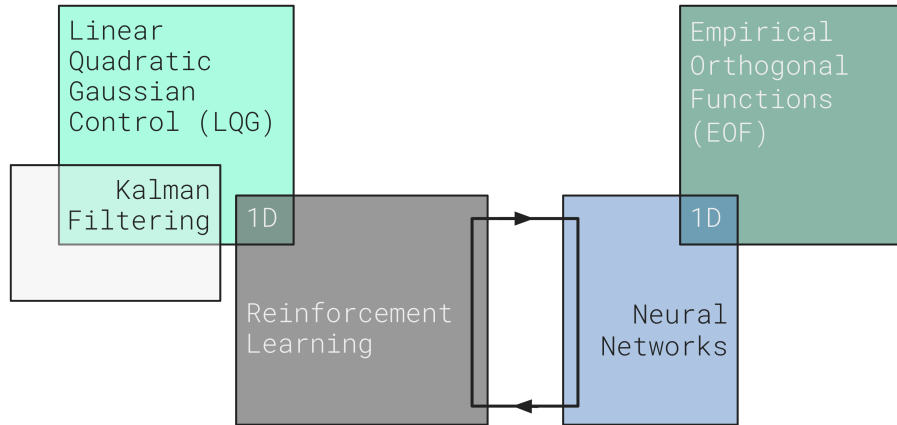
Figure 5.
Reinforcement learning is a particular way to frame a problem; a linear example of reinforcement learning is a linear quadratic Gaussian controller (which requires a Kalman filter as its estimator.) Neural networks are often used to solve reinforcement learning problems, and empirical orthogonal functions is equivalent to a neural net that has been constrained to a single layer and temporal evolution.

3. RESULTS FROM THE LITERATURE

3.1 Linear Methods

3.1.1 Empirical orthogonal functions

Arguably the first predictive controller for astronomical telescopes is from a work by Dessenne, 1998.[19] This uses a method quite similar to empirical orthogonal functions, except that it runs in closed-loop; instead of associating the full state of the wavefront in time it builds a similar history vector that includes both wavefront sensor residuals and DM commands from every iteration. The truth condition that the filter is optimized against uses the rejection transfer function (i.e., a high fidelity pseudo-open loop reconstruction) to reconstruct the full wavefront from the DM commands and wavefront sensor slopes. This method was implemented on-sky in 1999,[23] at the 1.52 meter visible telescope at Haute Provence. They cite a 30% relative Strehl increase over a classic integrator, which maps to an increase of $\sim 11\%$ to 14% Strehl over varying r_0 conditions at a central wavelength of 650 nm.

After Guyon proposed EOF in 2017,[14] it was implemented on the 8 meter Subaru/SCExAO in 2018.[18] An example of predictive control on/off shows visual improvement when comparing two on-sky images, and provides a factor of up to 3 improvement in contrast at close-in inner working angles.[24]

Shortly thereafter, EOF was also applied in simulation to the 8 meter VLT/SPHERE telemetry.[25] This application used telemetry data (in the form of wavefront sensor slopes and DM commands) to apply predictive control as if it were a post-processing method, and compare this to the actual on-sky performance of the integrator. van Kooten, 2020 found a 5.1 factor of improvement in RMS error over the residuals of the SAXO integrator, and estimated an additional 2.1 increase over an idealized AO system for VLT/SPHERE when applied to data taken at a central wavelength of 658 nm.

Finally, work started in 2019[15] to apply this on-sky at the 10 meter Keck observatory. Performance improvements have been shown with a similar post-processing telemetry application,[15, 17] and van Kooten, 2022[16] demonstrates an on-sky application. van Kooten, 2022 found a factor of 3 contrast improvement over the classic integrator at an inner working angle of 3-7 λ/D, and a 1.5 factor decrease in standard deviation when applied to L band coronagraphic observations, making the predictive controller consistently more robust than the integrator. A similar test also looked at Strehl from the Bracket γ narrowband filter, and found a Strehl increase from $\sim 19\%$ with the integrator to $\sim 23\%$ with the predictive EOF method.

3.1.2 Kalamn filtering

Haffert, 2021[26] takes a Kalman filtering-like approach, with a subspace predictive controller. Subspace predictive control builds and updates a predictive Kalman filter, with a state vector built of wavefront sensor slopes and DM commands; this work could also be considered an extension of the Dessenne, 1998[19] work, but without the need for rejection transfer function (or pseudo-open loop reconstructed) data to train against. This method, when applied to the MagAO-X daytime bench (an instrument for the 8 meter Magellan-Clay telescope) shows a contrast improvement by a factor of of up to 10-20 for low wind speeds and up to two orders of magnitude for higher wind speeds at a central wavelength of 760 nm.[26]

3.1.3 Linear quadratic Gaussian controllers

Linear quadratic Gaussian control was suggested as a controller for astronomical telescopes in 1993,[13] making 30 years of machine learning methods for wavefront control with the publication of this proceeding. The classic implementation of this method quickly gained popularity for the control of lower order modes. LQG runs as the standard low order controller for the VLT/SPHERE, where low order LQG performed with an RMS error of 0.87 mas on tip and tilt, as compared to a classic integrator at 12.8 mas.[27] The Gemini Planet Imager (GPI) also uses low order LQG control, in particular to control consistent and specific frequencies; for example one 27 Hz frequency induced a 19.5 mas tilt, and with the LQG 2.3 mas of tilt remained at a central wavelength of 1570 nm.[28]

The only application of this method as a full controller on-sky was on the 4.2 meter CANARY AO demonstrator at the William Herschell Telescope, which reports a Strehl of 50.9% for a classic integrator, and 61.9% for the LQG controller in the H band.[29] A reformulated data-driven LQG[30] was also tested on CANARY/WHT, finding up to 7% Strehl improvement (from 0.067 to 0.051 radians RMS) when run on-sky in the H band.

Predictive Fourier control (PFC)[31] is another LQG alternative, which learns wind speeds from telemetry data (often called a wind-driven predictor). Its original implementation was for GPI-like simulations, which showed an integrator residual RMS error of 37.2 nm, as compared to 21 nm for the predictive Fourier control.[31] Fowler, 2022 applied this method to Keck telemetry, finding an improvement in RMS error on by a factor of 2.5 when comparing PFC to on-sky integrator residuals.[17] Most recently lab results from the LLAMAS (Low Latency Adaptive Optical Mirror System) testbed show a decrease in temporal error power by up to 3, with varying wind speeds and spatial frequencies.[32] The predictive Fourier control method also allows for non-integer lags in the predictor, which makes for a highly tunable predictive corrector.

3.2 Neural Networks

3.2.1 Wavefront prediction

Some works have explored the use of neural networks (NN) for predicting the evolution of atmospheric turbulence or wavefront sensor telemetry.[21, 22, 33] These mainly focused on the prediction of (pseudo-)open loop measurements, and all show promising improvements over classical, non-predictive controllers (i.e., an integrator). We note that these methods build a nonlinear state-space model that predicts the evolution of turbulence, but do not build an optimal nonlinear controller; a classic integrator is often used as controller. From these works, Wong, 2021[21] used a simple feed-forward NN, while Hafeez and Liu used Long-Short Term Memory networks to deal with the temporal aspect of the data.[22, 33]

Swanson, 2022[34] used NN prediction in a closed-loop AO system, showing that in closed-loop there can be a mismatch between the training distribution and observed distribution, which can be mitigated by using an adverserial prior. Swanson, in particular, explored the role of guidestar magnitude and increased sky coverage, showing that the NN method was able to show major performance improvements for a 16th magnitude guidestar, from a Strehl ratio of 40% with an integrator to 70% with the NN, for a simulated 8 meter telescope in the R band.

3.2.2 Reinforcement learning

The challenge with nonlinear state-space models is that it is not trivial to find an analytical solution to the control problem with a quadratic cost function (unlike linear examples). To that end, reinforcement learning (RL) aims to derive optimal nonlinear controllers by e.g. using planning algorithms (model-based) or through direct optimization of a parameterized control law (model-free). An additional feature of RL algorithms is that one is not limited to a quadratic cost function of the state variables but is free to choose the optimization objective (i.e., reward), allowing one to directly optimize for e.g. Strehl or contrast.

Nousiainen, 2021[35] used model-based RL in a simulated 24x24 AO system with a Shack-Hartmann and demonstrated an improvement in Strehl and contrast compared to an integrator, showing that their algorithm is also able to deal with DM to wavefront sensor misregistrations, but notes that the algorithm may be too computationally expensive to run at kHz frequencies. A consecutive work (Nousiainen, 2022[36]) solved this issue by using a parameterized control law; they not only show significant contrast improvements in a simulated VLT-like system with a PWFS, and decreased speckle variance in the lab with MagAO-X, but also run a simulations of a 40 meter telescope in real time with this controller, increasing performance from 88% Strehl with a classic integrator to 90% Strehl for the RL method, calculated at 1650 nm. Most recently, Nousiainen demonstrated these results the GHOST testbench and showed promising initial results in a cascaded AO system.[37]

On the other hand, in the work of Landman, 2020, and 2021,[38,39] a model-free algorithm is chosen based on the Deep Deterministic Policy Gradient algorithm. Landman shows improvements in vibration mitigation in simulations and lab experiments and a large improvement in contrast in a simulated high-contrast imaging instrument on an 8 meter telescope (from 10^{-4} for the integrator to 2×10^{-5} for the RL method, calculated at 1000 nm). By utilizing a combination of an LSTM and CNN, the RL algorithm built a policy that adjusts to changing atmospheric conditions without online learning, making the controller highly adaptable to changing on-sky conditions.

Pou, 2022[40] also uses a model-free approach using the Soft Actor-Critic algorithm and considers a distributed approach using multi-agent RL. Pou shows improvements in Strehl for a simulated 8-meter telescope with Shack-Hartmann WFS and PWFS, from 85% Strehl to 91% Strehl calculated at 1650 nm.[41] However, for the Pyramid, this performance required a nonlinear wavefront reconstructor before the RL controller.

3.3 Are Non-Linear Methods Justified

Despite pointing to non-linearities as a potential goal for new wavefront control algorithms, some interesting work in the field has explored whether or not temporally evolving non-linear controllers are necessary. Multiple works,[22,21] have compared classic linear solvers (i.e., empirical orthogonal functions or a neural net with a single layer) to multi-layer approaches, only to find that when considering temporally evolving statistics the multi-layer systems converge to the same answers, sometimes with added noise in the multi-layer networks.

Haffert, 2021[26] argues that running in closed-loop maintains the condition of linearity for both Pyramid wavefront sensors (PWFS) and DM actuators, meaning non-linear reconstructions and system dynamics are not worth encapsulating in temporally evolving models. However, other works[36,39] demonstrate that a non-linear solver can learn the entire system model, including wavefront sensing response, deformable mirror dynamics, and DM to wavefront sensor misregistration, which leads to a more painless and flexible system implementation. Furthermore, Wong 2021[21] found that despite converging to the same solution, the neural network found the answer with ~5 seconds of training data, as compared to a linear predictor which needed ~40 seconds of training data.

4. CONCLUSION

With the 1993 Paschall & Anderson paper[13] on LQG controllers for astronomical telescopes, we celebrate 30 years of machine learning for wavefront control, over which machine learning approaches to wavefront control have been suggested, simulated, tested in the lab, and rarely tested on sky. A variety of methods show promising results, but many of these control algorithms go no further because they have yet to be consistently verified on-sky. Figure 6 compares performance from the described methods, either in reported Strehl ratio, or contrast.
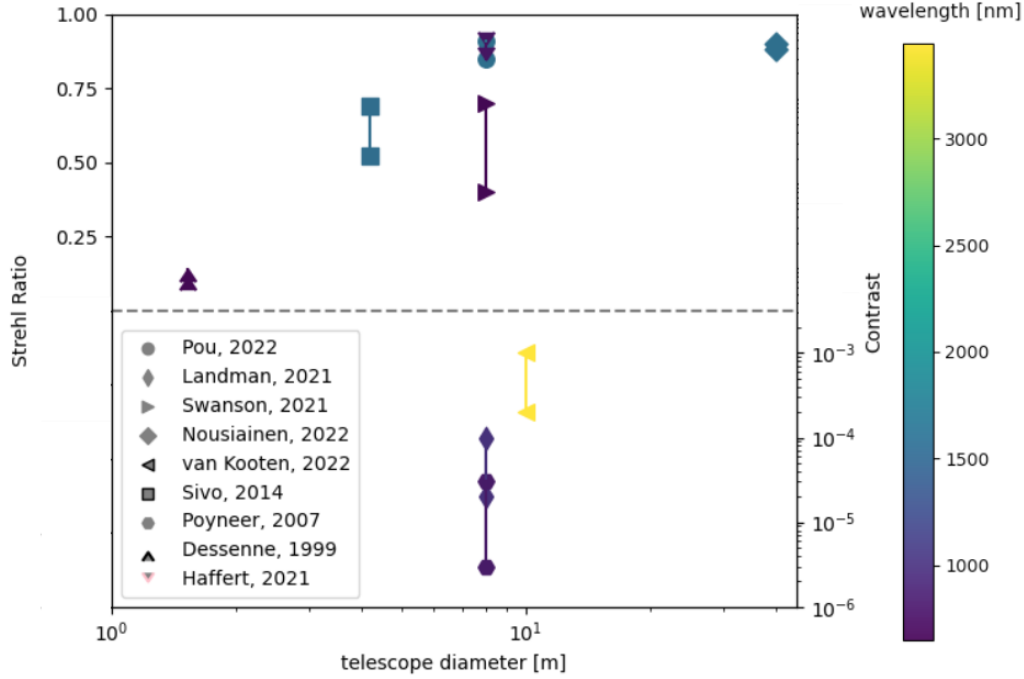
Figure 6.

Comparative performance of machine learning methods from the literature with increasing telescope diameter, connecting integrator performance to machine learning method performance. Everything above the dashed line is a reported Strehl ratio, and below a contrast. Color indicates wavelength. No outline: simulation, pink outline: lab results, and black outline: on-sky testing. Only van Kooten, 2022[16] was for a large telescope with a coronagraphic imager.

This is not a perfect comparison, as these performance reports span different wavelength regimes, actuator counts, wavefront sensors, and telescope diameters.

As we design the next generation of extremely large telescopes, only a few[16, 18] works show proven results on-sky (for a 8-10 meter class telescope controlling all modes with a ML method), making it difficult to promise or benchmark performance for future systems. Furthermore, most of these efforts focus on a single algorithm and its development and progression on a single system. Few efforts[17, 21, 22] have attempted to compare the performance of these methods side-by-side, to determine how these methods may improve performance, stability, or make for simpler problems across a uniform simulation, lab demonstration, or telescope system.

APPENDIX A. MATH MISTAKES IN FUNDAMENTAL PAPERS

A.1 Paschall & Anderson, 1993

The first paper to describe applying an LQG controller to an astronomical telescope system. The quadratic cost function described in Eq (17) of the original paper,[13] written as:

$$\mathbf{J} = \mathbf{E} \left\langle \mathbf{\Sigma}_{i=0}^{N} \frac{1}{2} \left[\mathbf{x^T}(t_i)\mathcal{X}\mathbf{x}(t_i + \mathbf{u^T}(t_i)\mathcal{U}\mathbf{u}(t_i) \right] \right\rangle \tag{9}$$

is missing a close paren, it should read:

$$\mathbf{J} = \mathbf{E} \left\langle \mathbf{\Sigma}_{i=0}^{N} \frac{1}{2} \left[\mathbf{x^T}(t_i)\mathcal{X}\mathbf{x}(t_i) + \mathbf{u^T}(t_i)\mathcal{U}\mathbf{u}(t_i) \right] \right\rangle \tag{10}$$

This is an arguably obvious mistake, as the vector $\mathbf{x}$ must apply to some discrete timestep.

A.2 Dessenne et al., 1998

Many credit Dessenne, 1998[19] as the first paper to describe a fully predictive controller. It builds a state vector ϕ, which holds both wavefront sensor residuals and deformable mirror commands. Equation (20) of the original work lays out the state vector and the predictive filter as:

$$\boldsymbol{\theta} = [b_1...b_{p-1}, a_0...a_{q-1}]^T$$
$$\phi(i) = [y(i-1)...y(i-p+1)]$$
$$(y - e'_{meas})(i-2)...(y - e'_{meas})(i-1-1)]$$

This expression has an extra bracket inserted. It should read:

$$\boldsymbol{\theta} = [b_1...b_{p-1}, a_0...a_{q-1}]^T$$
$$\phi(i) = [y(i-1)...y(i-p+1), (y - e'_{meas})(i-2)...(y - e'_{meas})(i-1-1)]$$

The entirety of the second two lines of the originaly Equation (20) make up a single expression for the state vector. This error has high potential for confusion as one could interpret it as a third expression, missing its definition, or two series of terms that should be multiplied as cross terms.

A.3 Poyneer et al., 2007

Poyneer, 2007[31] is one of two major predictors explored today. It makes use of a particular modal basis, complex Fourier modes. The expression for complex Fourier modes in the original work from Equation (1) is written:

$$f_{k,l}[x,y] = \cos\left(i2\pi\frac{kx+ly}{N}\right) + i\sin\left(i2\pi\frac{kx+ly}{N}\right) \tag{11}$$

which has two extra imaginary terms inserted. It should read:

$$f_{k,l}[x,y] = \cos\left(2\pi\frac{kx+ly}{N}\right) + i\sin\left(2\pi\frac{kx+ly}{N}\right) \tag{12}$$

Those familiar with complex modes would likely spot this mistake immediately, but due to the waning use of complex modes it may be easy to forget where the imaginary terms lie. It should also be noted that the cos and sin of imaginary numbers provides a real answer, so it is not immediately obvious in debugging results mathematically that this is incorrect.

REFERENCES

[1] Wong, A. P., Norris, B. R. M., Deo, V., Guyon, O., Tuthill, P. G., Lozi, J., Vievard, S., and Ahn, K., "Machine learning for wavefront sensing," in [*Adaptive Optics Systems VIII*], Schreiber, L., Schmidt, D., and Vernet, E., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **12185**, 121852I (Aug. 2022).

[2] Gendron, E. and Léna, P., "Astronomical adaptive optics. ii. experimental results of an optimized modal control.," *Astronomy and Astrophysics Supplement, v. 111, p. 153* **111**, 153 (1995).

[3] Males, J. R. and Guyon, O., "Ground-based adaptive optics coronagraphic performance under closed-loop predictive control," *Journal of Astronomical Telescopes, Instruments, and Systems* **4**(1), 019001–019001 (2018).

[4] Fowler, J., Haffert, S., van Kooten, M. A. M., Landmand, R. and, B. A., Hours, A., N'Diaye, M., Absil, O., Altinier, L., Baudoz, P., Belikov, R., Bonse, M. J., Bott, K., Brandl, B., Carlotti, A., Casewell, S., Choquet, E., Cowan, N. B., Desai, N., Doelman, D., Fogarty, K., Gebhard, T. B., Gutierrez, Y., Guyon, O., Herscovici-Schiller, O., , Juanola-Parramon, R., Kenworthy, M., Kleisioti, E., Konig, L., Krasteva, M., Laginja, I., Leboulleux, L., Mazoyer, J., Millar-Blanchaer, M. A., Mouillet, D., Por, E., Pueyo, L., Snik, F., van Dam, D., van Gorkom, K., and Vaughan, R. S., "Visible extreme adaptive optics on extremely large telescopes: Towards detecting oxygen in proxima centauri b and analogs," *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **in same proceedings** (2023).

[5] Kolmogorov, A., "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers," *Akademiia Nauk SSSR Doklady* **30**, 301–305 (Jan. 1941).

[6] Tatarskii, V. I., [*Wave Propagation in Turbulent Medium*] (1961).

[7] Taylor, G. I., "The Spectrum of Turbulence," *Proceedings of the Royal Society of London Series A* **164**, 476–490 (Feb. 1938).

[8] Poyneer, L., van Dam, M., and Véran, J.-P., "Experimental verification of the frozen flow atmospheric turbulence assumption with use of astronomical adaptive optics telemetry," *Journal of the Optical Society of America A* **26**, 833 (Mar. 2009).

[9] Dravins, D., Lindegren, L., Mezey, E., and Young, A. T., "Atmospheric Intensity Scintillation of Stars. III. Effects for Different Telescope Apertures," **110**, 610–633 (May 1998).

[10] Woolf, N., "Dome seeing.," **91**, 523–529 (Aug. 1979).

[11] Hastie, Trevor, T. R. and Friedman, J., [*The Elements of Statistical Learning, Data Mining, Inference, and Prediction, second edition*] (01 2017).

[12] Kalman, R. E., "A new approach to linear filtering and prediction problems," *Transactions of the ASME– Journal of Basic Engineering* **82**(Series D), 35–45 (1960).

[13] Paschall, R. N. and Anderson, D. J., "Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements," **32**, 6347–6358 (Nov. 1993).

[14] Guyon, O. and Males, J., "Adaptive Optics Predictive Control with Empirical Orthogonal Functions (EOFs)," *arXiv e-prints* , arXiv:1707.00570 (July 2017).

[15] Jensen-Clem, R., Bond, C. Z., Cetre, S., McEwen, E., Wizinowich, P., Ragland, S., Mawet, D., and Graham, J., "Demonstrating predictive wavefront control with the Keck II near-infrared pyramid wavefront sensor," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **11117**, 111170W (Sept. 2019).

[16] van Kooten, M. A. M., Jensen-Clem, R., Cetre, S., Ragland, S., Bond, C. Z., Fowler, J., and Wizinowich, P., "Predictive wavefront control on Keck II adaptive optics bench: on-sky coronagraphic results," *Journal of Astronomical Telescopes, Instruments, and Systems* **8**, 029006 (Apr. 2022).

[17] Fowler, J., Van Kooten, M. A. M., and Jensen-Clem, R., "Battle of the predictive wavefront controls: comparing data and model-driven predictive control for high contrast imaging," in [*Adaptive Optics Systems VIII*], Schreiber, L., Schmidt, D., and Vernet, E., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **12185**, 1218582 (Aug. 2022).

[18] Guyon, O., Sevin, A., Gratadour, D., Bernard, J., Ltaief, H., Sukkari, D., Cetre, S., Skaf, N., Lozi, J., Martinache, F., Clergeon, C., Norris, B., Wong, A., and Males, J., "The compute and control for adaptive optics (CACAO) real-time control software package," in [*Adaptive Optics Systems VI*], Close, L. M., Schreiber, L., and Schmidt, D., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **10703**, 107031E (July 2018).

[19] Dessenne, C., Madec, P.-Y., and Rousset, G., "Optimization of a Predictive Controller for Closed-Loop Adaptive Optics," **37**, 4623–4633 (July 1998).

[20] Recht, B., "A Tour of Reinforcement Learning: The View from Continuous Control," *arXiv e-prints* , arXiv:1806.09460 (June 2018).

[21] Wong, A. P., Norris, B. R. M., Tuthill, P. G., Scalzo, R., Lozi, J., Vievard, S., and Guyon, O., "Predictive control for adaptive optics using neural networks," *Journal of Astronomical Telescopes, Instruments, and Systems* **7**, 019001 (Jan. 2021).

[22] Hafeez, R., Archinuk, F., Fabbro, S., Teimoorinia, H., and Véran, J.-P., "Forecasting wavefront corrections in an adaptive optics system," *Journal of Astronomical Telescopes, Instruments, and Systems* **8**, 029003 (Apr. 2022).

[23] Dessenne, C., Madec, P. Y., and Rousset, G., "Sky implementationof modal predictive control in adaptive optics," *Optics Letters* **24**, 339–341 (Mar. 1999).

[24] Currie, T., Guyon, O., Lozi, J., Groff, T., Kasdin, N. J., Martinache, F., Brandt, T. D., Chilcote, J., Marois, C., Gerard, B., Jovanovic, N., and Vievard, S., "Performance and early science with the Subaru Coronagraphic Extreme Adaptive Optics project," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **11117**, 111170X (Sept. 2019).

[25] van Kooten, M. A. M., Doelman, N., and Kenworthy, M., "Robustness of prediction for extreme adaptive optics systems under various observing conditions. An analysis using VLT/SPHERE adaptive optics data," **636**, A81 (Apr. 2020).

[26] Haffert, S. Y., Males, J. R., Close, L. M., Van Gorkom, K., Long, J. D., Hedglen, A. D., Guyon, O., Schatz, L., Kautz, M., Lumbres, J., Rodack, A., Knight, J. M., Sun, H., and Fogarty, K., "Data-driven subspace predictive control of adaptive optics for high-contrast imaging," *Journal of Astronomical Telescopes, Instruments, and Systems* **7**, 029001 (Apr. 2021).

[27] Petit, C., Fusco, T., Sauvage, J. F., Sevin, A., Suarez, M., Charton, J., Baudoz, P., Beuzit, J. L., Buey, T., Dohlen, K., Feautrier, P., Fedrigo, E., Fleuzry, B., Gach, J. L., Hubin, N., Kasper, M., Mouillet, D., Perret, D., Puget, P., Sinquin, J. C., Soenke, C., and Wildi, F., "Integration, tests and laboratory performance of SAXO, the VLT-SPHERE extreme AO system," in [*Second International Conference on Adaptive Optics for Extremely Large Telescopes. Online at ¡A href="http://ao4elt2.lesia.obspm.fr"¿http://ao4elt2.lesia.obspm.fr¡/A*], 16 (Sept. 2011).

[28] Poyneer, L. A., Palmer, D. W., Macintosh, B., Savransky, D., Sadakuni, N., Thomas, S., Véran, J.-P., Follette, K. B., Greenbaum, A. Z., Mark Ammons, S., Bailey, V. P., Bauman, B., Cardwell, A., Dillon, D., Gavel, D., Hartung, M., Hibon, P., Perrin, M. D., Rantakyrö, F. T., Sivaramakrishnan, A., and Wang, J. J., "Performance of the Gemini Planet Imager's adaptive optics system," **55**, 323 (Jan. 2016).

[29] Sivo, G., Kulcsár, C., Conan, J.-M., Raynaud, H.-F., Gendron, É., Basden, A., Vidal, F., Morris, T., Meimon, S., Petit, C., Gratadour, D., Martin, O., Hubert, Z., Sevin, A., Perret, D., Chemla, F., Rousset, G., Dipper, N., Talbot, G., Younger, E., Myers, R., Henry, D., Todd, S., Atkinson, D., Dickson, C., and Longmore, A., "First on-sky SCAO validation of full LQG control with vibration mitigation on the CANARY pathfinder," *Optics Express* **22**, 23565 (Sept. 2014).

[30] Sinquin, B., Prengère, L., Kulcsár, C., Raynaud, H.-F., Gendron, E., Osborn, J., Basden, A., Conan, J.-M., Bharmal, N., Bardou, L., Staykov, L., Morris, T., Buey, T., Chemla, F., and Cohen, M., "On-sky results for adaptive optics control with data-driven models on low-order modes," **498**, 3228–3240 (Nov. 2020).

[31] Poyneer, L. A., Macintosh, B. A., and Véran, J.-P., "Fourier transform wavefront control with adaptive prediction of the atmosphere," *Journal of the Optical Society of America A* **24**, 2645 (Jan. 2007).

[32] Poyneer, L. A., Ammons, S. M., Kim, M. K., Bauman, B., Terrel-Perez, J., Lemmer, A. J., and Nguyen, J., "Laboratory demonstration of the prediction of wind-blown turbulence by adaptive optics at 8 kHz with use of LQG control," **62**, 1871 (Mar. 2023).

[33] Liu, X., Morris, T., Saunter, C., de Cos Juez, F. J., González-Gutiérrez, C., and Bardou, L., "Wavefront prediction using artificial neural networks for open-loop adaptive optics," **496**, 456–464 (July 2020).

[34] Swanson, R., Lamb, M., Correia, C. M., Sivanandam, S., and Kutulakos, K., "Closed loop predictive control of adaptive optics systems with convolutional neural networks," **503**, 2944–2954 (May 2021).

[35] Nousiainen, J., Rajani, C., Kasper, M., and Helin, T., "Adaptive optics control using model-based reinforcement learning," *Optics Express* **29**, 15327 (May 2021).

[36] Nousiainen, J., Rajani, C., Kasper, M., Helin, T., Haffert, S. Y., Vérinaud, C., Males, J. R., Van Gorkom, K., Close, L. M., Long, J. D., Hedglen, A. D., Guyon, O., Schatz, L., Kautz, M., Lumbres, J., Rodack, A., Knight, J. M., and Miller, K., "Toward on-sky adaptive optics control using reinforcement learning. Model-based policy optimization for adaptive optics," **664**, A71 (Aug. 2022).

[37] Nousiainen, J., Engler, B., Kasper, M., Helin, T., Heritier, C. T., and Rajani, C., "Advances in model-based reinforcement learning for adaptive optics control," in [*Adaptive Optics Systems VIII*], Schreiber, L., Schmidt, D., and Vernet, E., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **12185**, 121852O (Aug. 2022).

[38] Landman, R., Haffert, S. Y., Radhakrishnan, V. M., and Keller, C. U., "Self-optimizing adaptive optics control with reinforcement learning," in [*Adaptive Optics Systems VII*], Schreiber, L., Schmidt, D., and Vernet, E., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **11448**, 1144849 (Dec. 2020).

[39] Landman, R., Haffert, S. Y., Radhakrishnan, V. M., and Keller, C. U., "Self-optimizing adaptive optics control with reinforcement learning for high-contrast imaging," *Journal of Astronomical Telescopes, Instruments, and Systems* **7**, 039002 (July 2021).

[40] Pou, B., Ferreira, F., Quinones, E., Gratadour, D., and Martin, M., "Adaptive optics control with multi-agent model-free reinforcement learning," *Optics Express* **30**, 2991 (Jan. 2022).

[41] Pou, B., Smith, J., Quinones, E., Martin, M., and Gratadour, D., "Model-free reinforcement learning with a non-linear reconstructor for closed-loop adaptive optics control with a pyramid wavefront sensor," in [*Adaptive Optics Systems VIII*], Schreiber, L., Schmidt, D., and Vernet, E., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **12185**, 121852U (Aug. 2022).