# Optical Performance Perception Technology in the Embodied Intelligence-Based Optical Astronomical Telescopes

**Yiming Zhang**[a,b,c], **Zhixu Wu**[d,*], **Zhengyang Li**[a,b,*], **Chao Chen**[a,b,c], **Jianan Cong**[a,b,c], **Zijian Han**[a,b,c], **Tingting Liu**[a,b]

[a]Nanjing Institute of Astronomical Optics & Technology, Chinese Academy of Sciences, Nanjing 210042, China
[b]CAS Key Laboratory of Astronomical Optics & Technology, Nanjing Institute of Astronomical Optics & Technology, Nanjing 210042,China

[c]University of Chinese Academy of Sciences, Beijing 100049, China
[d]Institute of Space Science and Technology, Nanchang University, 999 Xuefu Avenue, Nanchang 330031, China

**Abstract.** Embodied Intelligence-based optical astronomical telescopes represent a AI-driven approach to autonomous perception, decision-making, and optimization of observational performance under dynamic environmental conditions, will provide intelligent operation to the widely distributed unmanned astronomical telescopes. The perception of astronomical telescope observation performance, primarily manifested in the optical performance perception, can be achieved through multi-field curvature wavefront sensors, wavefront slope sensors, and image quality sensors. These devices measure physical quantities represented as point spread function (PSF) or wavefront aberration data. By utilizing the perception data of optical performance, combined with end-to-end imaging simulation and digital twin technology, telescopes can achieve active optical correction and active alignment, intelligently maintaining and optimizing optical imaging quality. With the further integration of multi-modal sensing, machine learning, and large language model, embodied intelligence will drive the transition of telescopes from ”precision instruments” to ”intelligent instruments,” providing new technological pathways for unmanned astronomical observatories.

**Keywords:** Embodied Intelligence; Active Optics; Multi-model Sensing; Autonomous Observation.

*Zhengyang Li, zyli@niaot.ac.cn *Zhixu Wu, wuzhixu@ncu.edu.cn

## 1 Introduction

Modern large astronomical telescopes are highly complex “optical-mechanical-electrical-software” integrated systems, with their system complexity primarily manifested in the coordination of multiple instruments, precise pointing and tracking, and the use of active optics and adaptive optics.1 Large telescopes typically have multiple focal points, and different focal points require the mounting of scientific instruments such as spectrometers and scientific CCD cameras, with complex switching and coordination;2 Achieving high-precision pointing and tracking accuracy in such massive structures relies on a large number of position encoders, gyroscopes, and complex control systems;3 Active optics compensate for mirror deformation to maintain high imaging quality, requiring hundreds to thousands of force actuators, position sensors, complex support structures,

AOPC 2025: Astronomical Technologies and Instrumentation,
edited by Jian Ge, Proc. of SPIE Vol. 13964, 139640A
 · 0277-786X · doi: 10.1117/12.3082938

real-time mechanical model calculations, and closed-loop control loops;[4] Adaptive optics compensate for rapid high-order aberrations caused by atmospheric turbulence in real time (millisecond level), requiring equipment including wavefront sensors, deformable mirrors with hundreds to thousands of force actuators, real-time wavefront reconstruction and control algorithm calculation units, and guide star systems.[5] These systems integrate a massive amount of heterogeneous sensors, actuators, control loops, and real-time computing units, forming an extensive and highly coupled system.

Driven by the urgent need for high-quality imaging from modern large-scale astronomical telescopes and the growing demand for autonomous observations (unmanned, real-time event response), the embodied intelligence of telescopes has become a core development direction for modern telescope design and operation.[6] The goal of embodied intelligent telescopes is to achieve intelligent maintenance and optimization of imaging quality based on optical performance perception, as well as unattended autonomous observation, through perception networks, intelligent decision-making, and high-speed execution units,[7] thereby driving the transition of telescopes from ”precision instruments” to ”intelligent instruments.” Among these, the perception of the observational performance of embodied intelligent telescopes is primarily manifested in the perception of the telescope’s optical performance. Currently, some exploratory research has been conducted on the intelligent perception of telescope observational performance. For example, Crenshaw et al. proposed a deep learning model for wavefront estimation in the Rubin Observatory’s active optical system. Comparisons with Batoid and PhoSim simulations showed that this model is 40 times faster than the traditional transport intensity equation (TIE) baseline algorithm, with error reductions of 2 times, 5 times, and 14 times under ideal conditions, vignetting, and blending scenarios, respectively, significantly improving optical quality and expanding the effective survey area.[8] Jia

et al. introduced a denoising autoencoder (DAE) for PSF modeling of large-field-of-view small-aperture telescopes. They added noise/aberrations to calibrated data or simulation templates to generate training samples, and then used a convolutional neural network to learn the PSF manifold space. This method overcomes the limitations of principal component analysis (PCA) under low signal-to-noise ratios and insufficient spatial sampling, enhancing robustness against vignetting and blending.9 Peterson et al. proposed the Photon Simulator (PhoSim), a comprehensive simulation framework based on photon Monte Carlo methods for generating high-fidelity astronomical images from optical survey telescopes; This framework fully simulates the entire chain of effects from source photon sampling to electron readout through layered atmospheric phase screens, mixed geometry/Fourier optical propagation, and device physics models, supporting large-scale simulations for telescopes such as LSST.10 Wu et al. developed a machine learning-based alignment method using star image PSF ellipticity distributions for active alignment in wide-field telescopes; by learning the mapping of misalignment parameters and field-dependent PSF patterns through a two-step neural network (coarse tuning and fine tuning), achieving secondary mirror position error $< 5$ $\mu$ m and tilt angle error $< 5$" in Mephisto telescope simulations, enabling high-resolution imaging during observations. Compared to traditional wavefront sensing methods, this approach significantly simplifies system structure and improves real-time performance.11 These intelligent and imaging simulation methods will help modern telescopes achieve intelligent full-process capabilities, from optical performance sensing and telescope state inference to active optical correction.

Unlike fully data-driven embodied intelligence, this paper proposes a data- and physics-driven embodied intelligent telescope optical performance perception technology in China. By integrating data-driven and physics-constrained approaches, this technology retains the flexibility of data-

driven methods while enhancing the interpretability and reliability of models by applying physical laws. First, by constructing a full-process imaging simulation of the image quality of an embodied intelligent telescope and a data twin, the image quality degradation process of the embodied intelligent telescope is understood, and driving data is generated. Then, by measuring physical quantities such as the point spread function or wavefront aberration data using multi-field curvature wavefront sensors, wavefront slope sensors, and image quality sensors, the imaging state of the telescope is intelligently perceived. Finally, by applying the perceived optical performance data, combined with the full-process imaging simulation and data twin technology, the model achieves active optical correction and active alignment of the telescope, intelligently maintaining and optimizing optical imaging quality.

## 2 Overall framework of the embodied intelligent telescope

As shown in Figure 1, the proposed intelligent telescope system in this study employs a three-stage progressive state evolution framework (left) and a multimodal intelligent closed-loop system (right) within a collaborative architecture. The embodied intelligent telescope consists of a classic three-layer architecture, known as "perception-inference-execution," comprising the perception layer, inference layer, and execution layer, respectively. The perception layer integrates data streams from environmental sensors and wavefront sensors,[12] such as temperature and wavefront; the inference layer applies pre-trained theoretical models, multimodal models, and large models to diagnose the root causes of state anomalies in real time, such as secondary mirror misalignment and temperature deformation; the execution layer outputs correction commands to the force actuator array (five degrees of freedom pose compensation of the secondary mirror and deformation compensation of the primary mirror).[13]

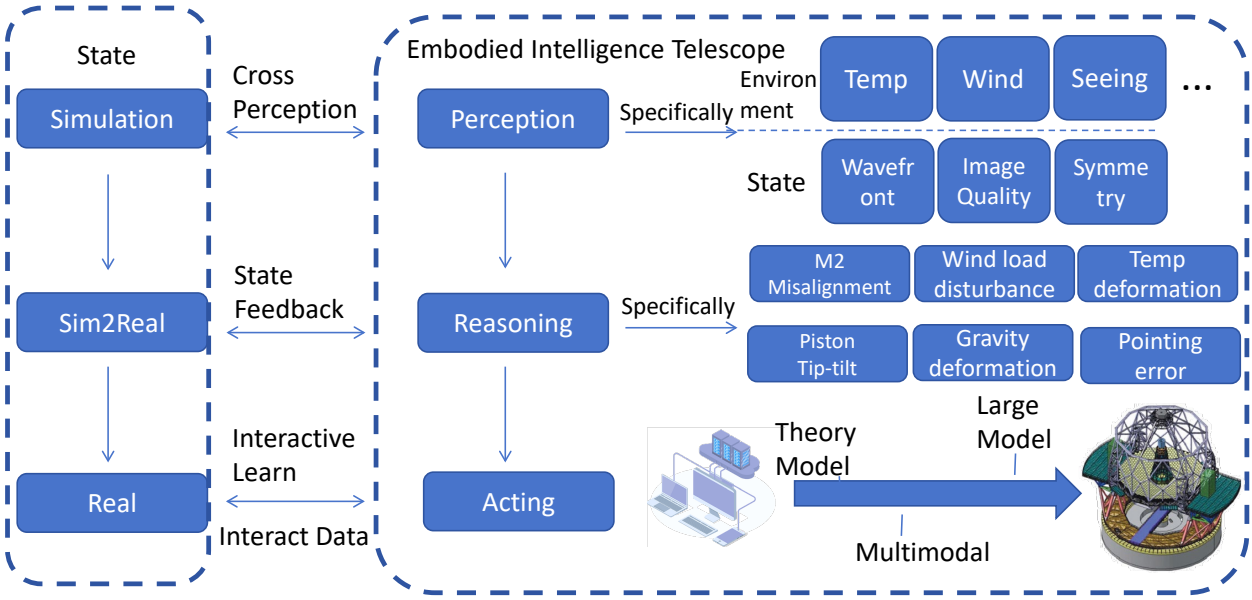


**Fig 1** Framework diagram of an embodied intelligent telescope. The left side illustrates the evolution of status, from analog to sim2real to real; the right side displays the embodied intelligent telescope framework, divided into three layers: perception, reasoning, and execution. The perception status includes temperature and wavefront, while the reasoning results include misalignment and gravity deformation. The model is divided into three stages: theory, multimodal, and large model.

The state evolution of embodied intelligent telescopes can be divided into three stages:

1. Theoretical model-driven stage (simulation state interaction)

In the early stages of system development, to address the bottleneck of limited real-world observational data, an independent theoretical model cluster was constructed. Through full-process imaging simulation (e.g., atmospheric disturbances, mirror deformation, tracking errors), multidimensional simulation state data were generated, and specialized neural network simulation models tailored for each subsystem (pointing tracking, active optics, adaptive optics, etc.) were established.14 At this stage, state interactions depend on the simulation environment, laying the theoretical foundation for subsequent real-world observations.

2. Multimodal model replacement phase (simulation-real state fusion interaction)

By deeply integrating telescope observation data sets, multimodal models are gradually intro-

duced to replace theoretical models.15 The specific implementation includes: (1) training neural networks based on real-state data; (2) constructing a hybrid state interaction layer: theoretical models and multimodal models run in parallel, and decision fusion is achieved through confidence weighting;(3) Switching mechanism: when the coverage of real data for a particular state reaches the threshold, the real data for that state can be used together with other real data to train the multi-modal model, which can then replace the simulation model for that state. At this stage, the system exhibits a hybrid simulation-real state characteristic. Through digital twin technology (processing simulated data and real data together), the efficiency, deviation, and defects of the theoretical model are quantified and analyzed to achieve a gradual enhancement of environmental adaptability.

3. Multimodal model dominance phase (real state interaction)

When deep integration based on real-time observation data is achieved, a multi-modal model training process that integrates real-time data is executed to realize an end-to-end perception-decision loop: control commands for actuators are generated directly from raw sensor data. Ultimately, a pure, real-time, state-driven, intelligent observation loop is achieved, significantly improving autonomous response capabilities in complex astronomical environments.

The development of embodied intelligent telescopes is a gradual process. Based on the data-driven characteristics of artificial intelligence models, with the continuous accumulation of embodied intelligent telescope datasets and the dimensional completeness of observation modalities, the telescope's perception state evolves from a simulated state to a real state, the telescope's decision-making model transitions from a single modality to a multi-modality, and the theoretical model evolves toward a large-scale model. A unified, large-scale model architecture achieves cognitive dimensionality enhancement by integrating the achievements of the first two stages. It inherits the physical constraints and interpretability of theoretical models, deepens the anomaly coupling

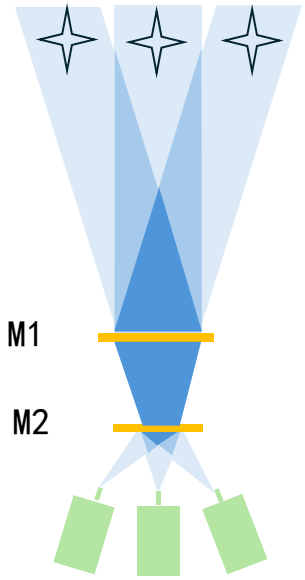


**Fig 2** Schematic diagram of end-to-end full-process imaging simulation (including atmosphere, optical system, and detector).

reasoning capabilities of multi-modal models, and directly generates control commands from raw sensor streams. This enables intelligent decision-making and execution for telescope anomalies, achieving self-iterative upgrading of embodied intelligent telescopes.

## 3 Telescope optical performance perception based on physical and data-driven coordination

### *3.1 Image quality throughout the entire imaging process simulation and data twins*

End-to-end imaging simulation refers to the reproduction of the entire physical process from photon generation to final electron readout, covering the complete chain from the atmosphere to telescope optics and detectors (Figure 2). The goal is to generate images consistent with actual observations and accurately reproduce all measurable attributes (such as point spread function morphology and background intensity levels).16

The simulation process is implemented using the photon Monte Carlo method 17, a high-fidelity light transmission simulation technology based on probabilistic sampling. Its core idea is to reproduce the physical interactions between light and matter by tracking the random propagation paths of a large number of independent photons. The key technical steps include photon generation, atmospheric transmission, optical systems, and detector path propagation. Photon generation is the initialization process of astrophysical photons, involving random sampling of position, time, wavelength, and morphology.18 The atmospheric module simulates the interaction between photons and turbulence, as well as molecular absorption. The turbulent phase screen model is used to simulate the random perturbations of light waves caused by atmospheric turbulence.19 Optical modules handle the reflection/refraction of photons with mirrors and lenses. The core is iterative ray tracing. Light intersection calculation is a core component of optical system simulation.20 The detector module includes photoconversion, charge diffusion, and readout. The key points are the quantum efficiency model and electric field-dependent diffusion.21

Twin simulation technology generates simulation data through full-process imaging simulation and compares it with real data, converting the physical effects of instruments and the environment into quantifiable system errors. The process is illustrated in Figure 3, which verifies the feasibility and robustness of existing theoretical models and optimizes and improves them.22

Full-process imaging simulation and data twin technology play a crucial role in enhancing the perception of image quality capabilities of telescopes. This provides a reliable data source and verification and optimization means for establishing optical performance perception models for telescopes.

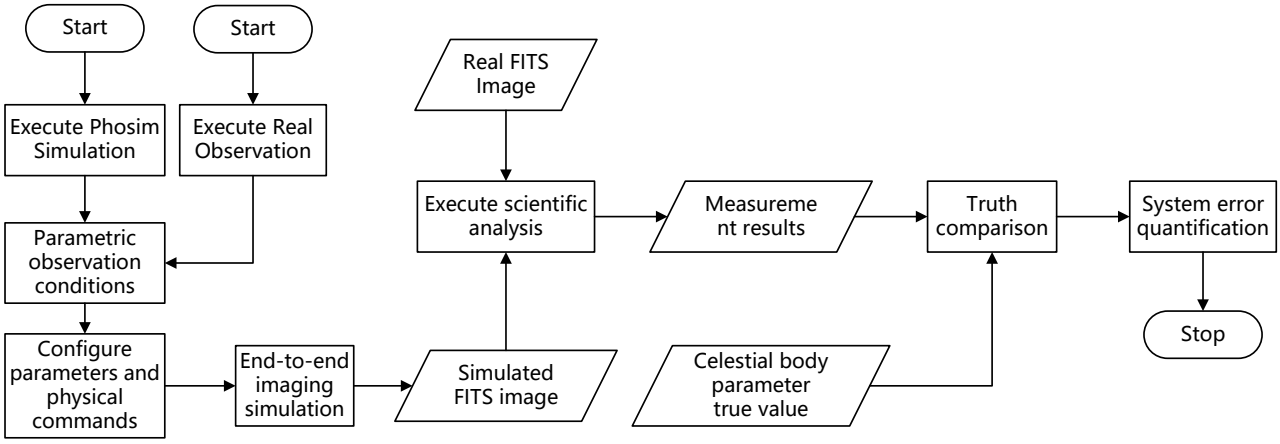


**Fig 3** Data twin simulation flowchart. Based on end-to-end imaging simulation, real FITS images are scientifically analyzed and compared with simulated FITS images to quantify system errors.

### *3.2 Image quality sensing*

In 2022, the research team explored an evaluation method for scientific observation images that can efficiently account for various types of collimation errors.11 First, based on the KSB (Kaiser-Squires-Broadhust) model describing galaxy structures, three parameters of star image quality under no external interference were described. Then, based on this description, it was found that the description of scientific observation image quality corresponding to collimation misalignment errors is not limited to focal plane tilt and eccentricity degrees of freedom but can also correspond to the eccentricity and tilt errors of the secondary mirror (as shown in Figure 4). The following conclusions are drawn: (a) Optical misalignment disrupts the point-symmetry property of the ideal optical system along the e1 and e2 axes; (b) The ellipticity e1 and e2 vary continuously with the field of view; (c) The ellipticity values vary nonlinearly with the misalignment error.

The KSB (Kaiser-Squire-Broadhurst) model is a core method used in the field of weak gravitational lensing to quantify the shape distortion of galaxies. It detects the distribution of matter

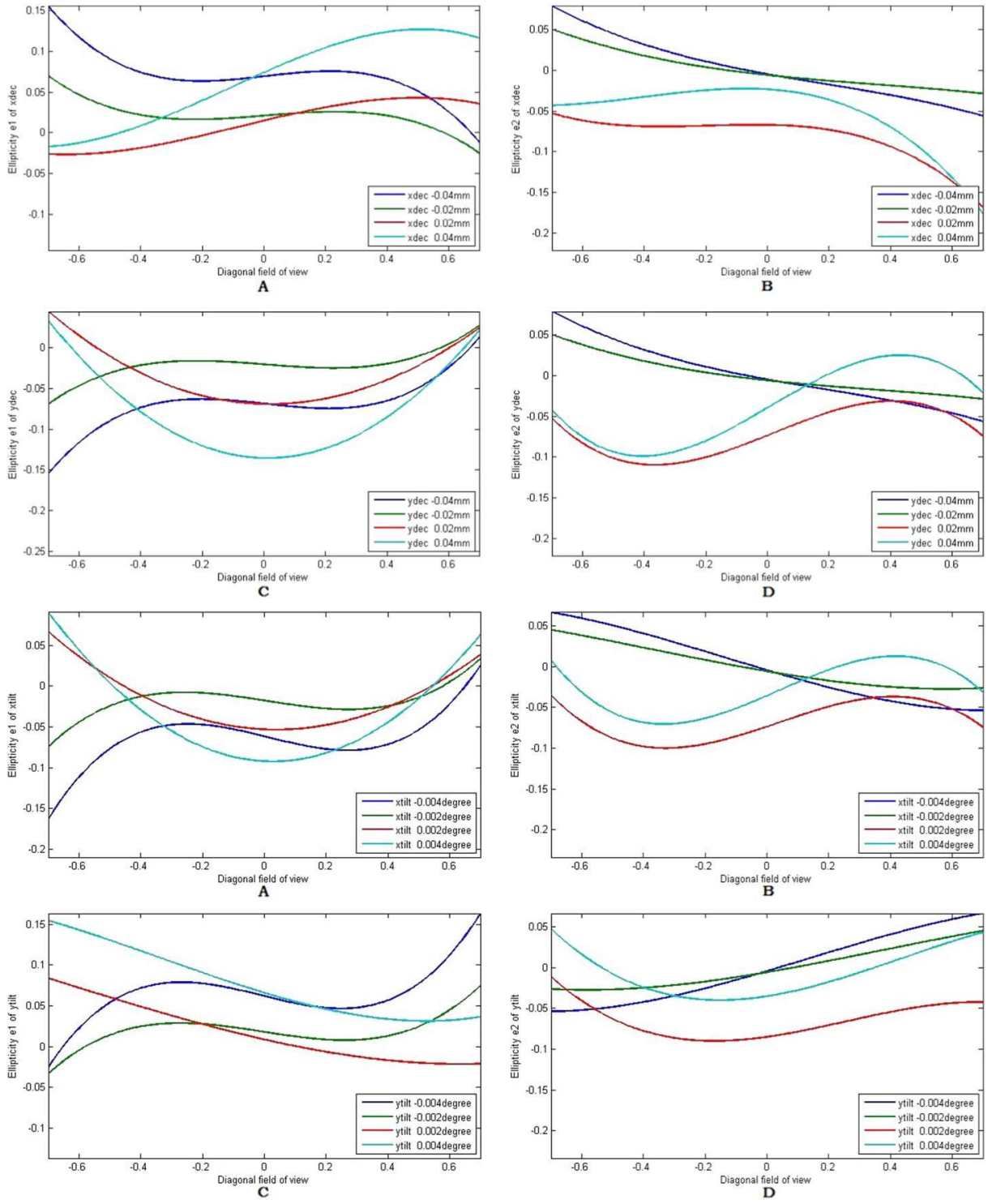


**Fig 4** The upper part indicates PSF ellipticity e1 and e2 with decentering (X, Y) of the secondary mirror, (A), (B). The distribution patterns of e1 and e2 when the secondary mirror of the optical system has x-decent misalignment. (C), (D) The e1 and e2 patterns of the secondary mirror with y-decent. The lower part indicates PSF ellipticity e1 and e2 with tilt (X, Y) of the secondary mirror. (A) Moreover, (B) displays the distribution patterns of e1 and e2 when the secondary mirror of the optical system has x-tilt misalignment; (C) and (D) are the e1 and e2 patterns of the secondary mirror with y-tilt.

(including dark matter) in the universe by parameterizing the ellipticity, radius, and azimuth of celestial objects.23 Its physical significance is described as follows:

$$
\begin{aligned}
Q_{i,j} &= \iint_{(H_x,H_y)} weight\,(H_x,H_y)\,PSF\,(\delta_1,\delta_2,\ldots,\delta_i,H_x,H_y)\,H_xH_ydH_xdH_y \\
e_\alpha &= \frac{Q_\alpha}{T}
\end{aligned}
\tag{1}
$$

Where,$\alpha = 1 or 2, H_1 = H_x - H_{xcenter}, H_2 = H_y - H_{ycenter}, Q_1 = Q_{11} - Q_{22}, T = Q_{11} + Q_{22}$, H is the field of view.

In practice, stellar point-mass degradation arises from nonlinear coupling between multiple sources of interference. Using star image quality to describe the collimation state of a misaligned optical system essentially involves estimating aberrations using the PSF generated by the telescope optical system. However, ground-based telescope star images are affected by optical-mechanical structural diffraction effects, atmospheric jitter, telescope tracking errors, differences in star image brightness, and imaging device sampling rates, and these interference factors exhibit complex nonlinear coupling relationships. Figure 5 shows the simulation of PSF image quality degradation as various interference factors are gradually added in Mephisto.16 Where, (a) is the optical model imaging; (b) is the addition of primary mirror disturbance; (c) is the addition of secondary mirror disturbance; (d) is the addition of detector disturbance; (e) is the addition of dome seeing effect; (f) is the addition of atmospheric turbulence effect. To achieve precise identification of telescope status using star image measurements, it is necessary to study new technologies for efficiently removing other interference sources that affect star image quality, identifying static PSF introduced by the telescope, and thereby separately describing changes in star image quality caused by misalignment.

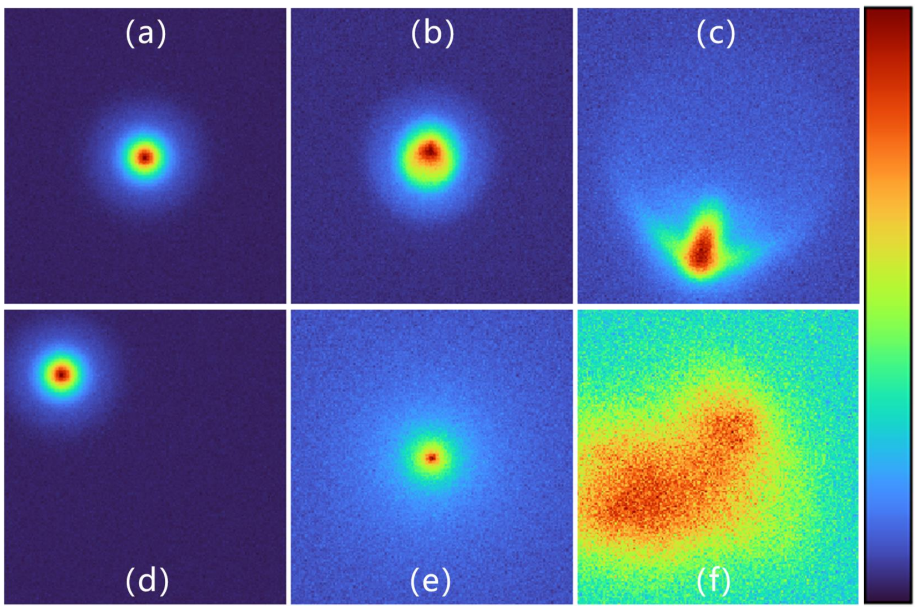


**Fig 5** Image quality degradation results after gradually introducing different disturbance factors. (a) Optical model imaging; (b) introduction of primary mirror disturbance; (c) introduction of secondary mirror disturbance; (d) introduction of detector disturbance; (e) introduction of dome seeing effect; (f) introduction of atmospheric turbulence effect.

In 2020, Jia Peng et al. proposed a data-driven modeling method based on denoising autoencoders (DAE) to address the challenge of modeling PSF spatiotemporal variations caused by off-axis aberrations, low signal-to-noise ratio (S/N), and insufficient spatial sampling rates in wide-field small-aperture telescopes (WFSATs).9 This method utilizes PSF templates generated from real observations or numerical simulations, constructs a training set by adding random noise and aberrations, and enables the DAE to learn the manifold space of the PSF. This method overcomes the limitations of traditional analytical models (such as the Moffat function) regarding symmetry assumptions. It addresses the bias issues of PCA methods in low S/N samples, laying a technical foundation for monitoring the status of optical systems and post-processing data.

Figure 6 shows the original image in the first row, the image with added noise $\sigma$=0.003 in the second row, and the PSF image reconstructed using the DAE PSF model in the third row. The image quality of scientific observations from ground-based telescopes is affected by the combined

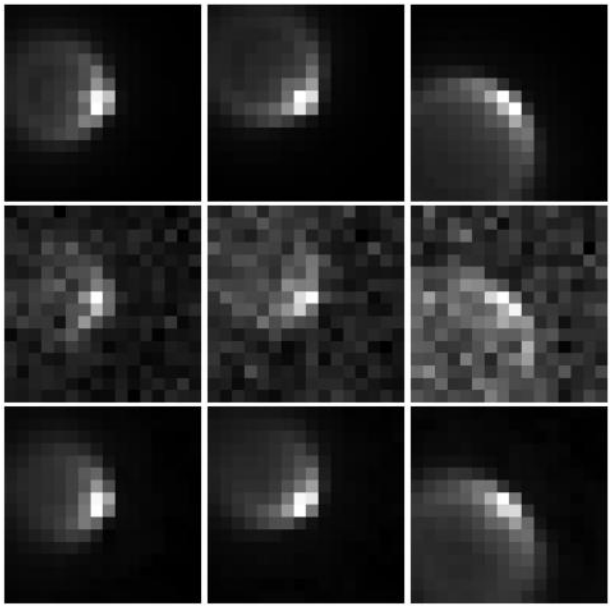

**Fig 6** Comparison of DAE PSF model reconstruction results. The first row shows the original image, the second row shows the image with noise added ($\sigma$=0.003), and the third row shows the PSF image reconstructed using the DAE PSF model.

effects of various disturbances, resulting in complex spatiotemporal degradation of the point spread function (PSF). Traditional methods struggle to decouple these factors accurately. The technique of using PSF for image quality sensing quantifies core parameters such as star image ellipticity (e), half-peak full width (R), and azimuth ($\theta$), establishing a direct correlation between scientific observation images and collimation error. This provides a physical basis for high-precision identification of telescope collimation status. Combined with data-driven methods, the PSF sensing model can robustly reconstruct PSF structure under low signal-to-noise ratio (S/N) and strong interference conditions, laying the foundation for achieving closed-loop control of optical system misalignment errors.

### *3.3 Multi-field curvature wavefront sensor*

Curvature wavefront sensing reconstructs the wavefront phase by measuring the difference in light intensity distribution (i.e., the curvature signal) between the defocused planes (in front of and be-

hind the focal point). The curvature signal is directly related to the Laplace operator and boundary slope of the wavefront.[24] Curvature wavefront sensing is based on the intensity transfer equation (ITE), which describes the relationship between the change in light intensity along the optical axis and the wavefront phase:

$$-k\frac{\partial I}{\partial z} = \nabla \cdot (I\nabla\phi) \tag{2}$$

Where, I is the light intensity distribution,$\phi$ is the wavefront phase, z is the optical axis direction, and $k = 2\pi/\lambda$ is the wave number.

Under uniform light intensity ($I \approx I_0$) and the para-axis approximation, ITE is simplified to the Poisson equation:

$$-\frac{k}{I_0}\frac{\partial I}{\partial z} = \nabla^2\phi - \delta_c\frac{\partial\phi}{\partial n} \tag{3}$$

Where, $\nabla^2\phi$ is the wavefront curvature (Laplace term), and $\delta_c\frac{\partial\phi}{\partial n}$ is the boundary normal slope (Dirac delta function term).

The curvature signal S is calculated using defocused images before ($I^+$) and after ($I^-$), which is expressed as a Poisson equation with Neumann boundary conditions:

$$S(r) = \frac{-1}{\Delta z}\frac{I_+(r) - I_-(r)}{(I_+(r) + I_-(r))} \approx \nabla^2\phi - \delta_c\frac{\partial\phi}{\partial n} \tag{4}$$

Roddier et al. proposed an iterative Fourier transform method in 1991 to solve the difficulty of boundary treatment. In principle, this method applies to wavefront restoration with any boundary shape.[25]

FFT iterative solution:

$$\phi_{\text{est}} = \mathcal{F}^{-1}\left[\frac{\mathcal{F}(S)}{-4\pi^2(u^2+v^2)}\right] \tag{5}$$

Where, $\mathcal{F}$ is the fast Fourier transform, $\mathcal{F}^{-1}$ is the inverse fast Fourier transform, and u and v are frequency domain coordinates.

The fast focal ratio system is susceptible to mirror deformation and component misalignment. Single-field measurements cannot capture the spatial variability of aberrations across the entire field of view, so it is necessary to measure wavefront aberrations in multiple fields of view. Liang Ming et al. tested the stability of the sensitivity matrix solution for multi-field curvature wavefront sensing.[26] Multi-field wavefront sensing enhances the stability of the misalignment solution. Additionally, deploying multiple wavefront sensors improves the probability of locating the signal source. As shown in Figure 7, Mephisto is equipped with four-field curvature wavefront sensors, providing robust support for acquiring wavefront aberration data.

Curvature wavefront sensing achieves high-precision reconstruction of wavefront aberrations by measuring changes in the intensity distribution of light at the defocus plane, utilizing a lens-free, vibration-resistant hardware structure. This technology has made a breakthrough in solving the real-time aberration detection challenge in active optics and adaptive optics for large-aperture telescopes. Its advantages include an extensive dynamic range ($> 1\lambda$ RMS aberration), a weak light detection capability (the multi-field Split CCD design fully utilizes edge field starlight), and resistance to environmental interference (frequency domain filtering suppresses noise), providing a critical foundation for closed-loop correction to enhance telescope imaging quality.

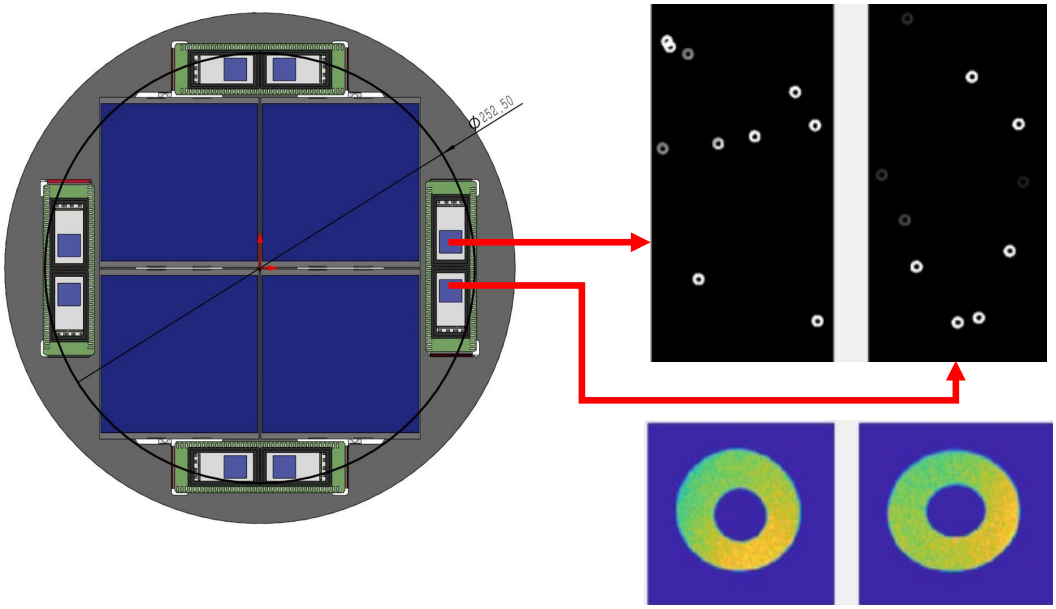


**Fig 7** Mephisto is equipped with four-edge field curvature wavefront sensors, each consisting of two parts, one for pre-focus imaging and one for post-focus imaging. On the left is Mephisto's focal plane; on the right above are the two fields of view of the cut CCD, and below is the wavefront reconstructed from defocused images.

## 4 Data-driven telescope status inference

The aberrations caused by misalignment in binocular or trinocular telescopes have been widely confirmed. Baranne and Wetherell derived the coma aberration formula for non-coaxial binocular systems in 1972.27 For more general cases, Su Dingqiang analyzed the coma aberration when the secondary mirror is eccentrically offset (decentered) and tilted (tip-tilt), and the primary and secondary mirror axes are non-coplanar, proving that coma aberration can be treated analogously to vector operations.28 Based on the vector aberration theory (Nodal Aberration Theory) proposed by Shack and Thompson, orthogonal Zernike polynomials have been widely used for the numerical expression of full-field distributed aberrations.29 Algorithms such as inverse optimization, damped least squares, and principal component analysis can achieve high-precision misalignment analysis.30

However, obtaining full-field aberration distribution during telescope observations is highly

challenging and relies on complex wavefront sensing systems. Four-edge field curvature wavefront sensors are commonly used for large-field imaging (e.g., JST/T250 and LSST).31 However, such systems require simultaneous measurement of pre-focus and post-focus images, and a fast focal ratio can easily cause vignetting. Additionally, curvature wavefront sensors must compensate for inherent aberrations across different fields of view (FOV) and rely on expensive CCD stitching technology, significantly increasing camera costs.

The impact of optical system misalignment on star image quality degradation involves two aspects: first, for a single field of view, multi-degree-of-freedom optical misalignment affects the diffusion morphology of the star point spread; second, for the entire field of view, multi-degree-of-freedom optical misalignment affects the distribution pattern of the star point spread across the entire focal plane. For example, Figure 8 shows the results of star image measurements after manually introducing optical system misalignment. It can be observed that the diffusion patterns of star point spreads vary significantly. In scientific observation images, the distribution of star images exhibits a random, discrete sampling distribution, and the overall field-of-view distribution pattern is complex to describe precisely. Since there are no comprehensive numerical description methods for star point spread diffusion patterns and the overall field-of-view distribution pattern, the underlying mechanism linking optical system misalignment and star image quality degradation remains unclear.

Unlike wavefront error detection, the ellipticity distribution pattern of the point spread function (PSF) can also characterize system misalignment 32, and PSF measurement based on scientific images is more convenient.33 Wu et al. proposed a machine learning-based alignment method that directly calculates misalignment parameters using star images from a wide-field telescope. We use a complex R-C system (Mephisto) as a model and apply multiple sets of perturbations to the

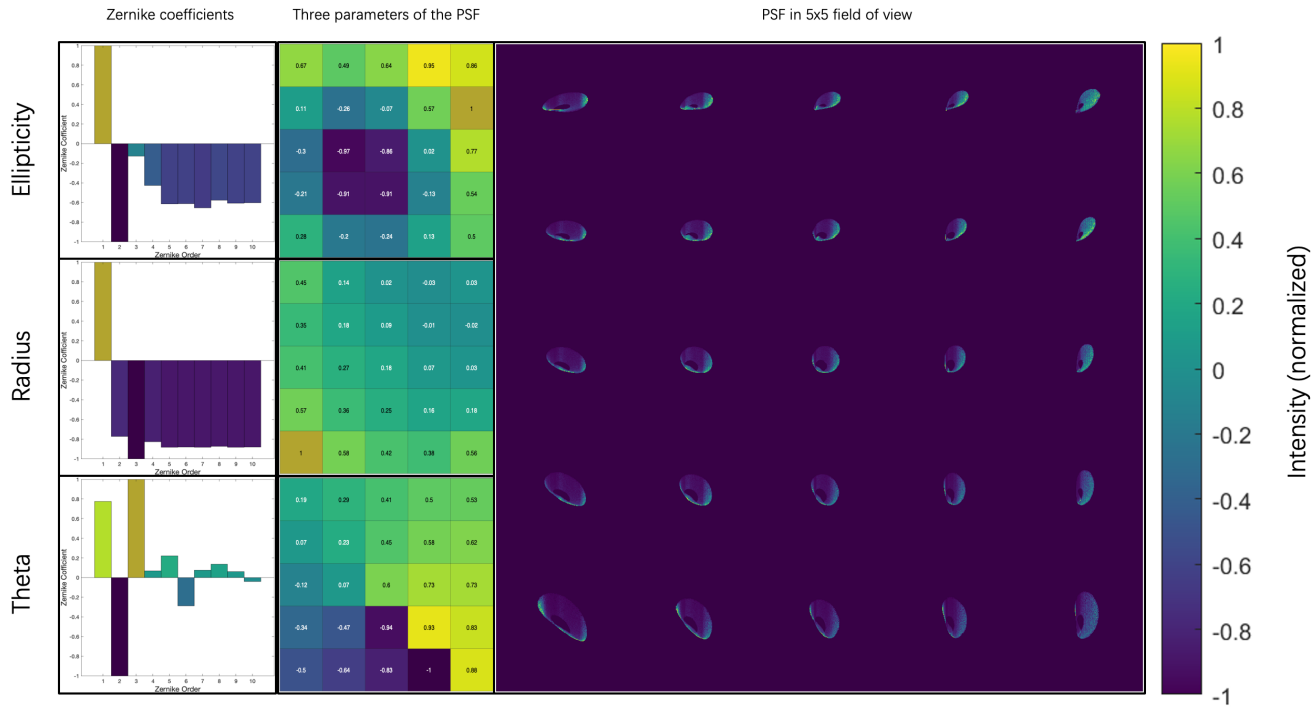


**Fig 8** Measurement results of the full field distribution of star images after introducing optical system misalignment. The right column is the PSF images of 25 FOVs; the middle column shows the calculated three parameters of the 25 PSFs, and the three parameters are ellipticity, radius, and theta; the left column is the fitted Zernike coefficients of the three parameters in the middle column.

full-field PSF distribution. Since it is difficult to establish an explicit mathematical model linking misalignment parameters to the full-field PSF distribution, machine learning can automatically construct analytical models.34 The inherent flexibility and adaptability of nonlinear neural networks (Figure 9) make them an ideal choice for addressing this problem.

## 5 Conclusions

This paper proposes a physically and data-driven embodied intelligent telescope optical performance perception technology, which integrates full-process imaging simulation, multimodal sensing, and machine learning methods to construct a closed-loop framework from optical performance perception to active correction.

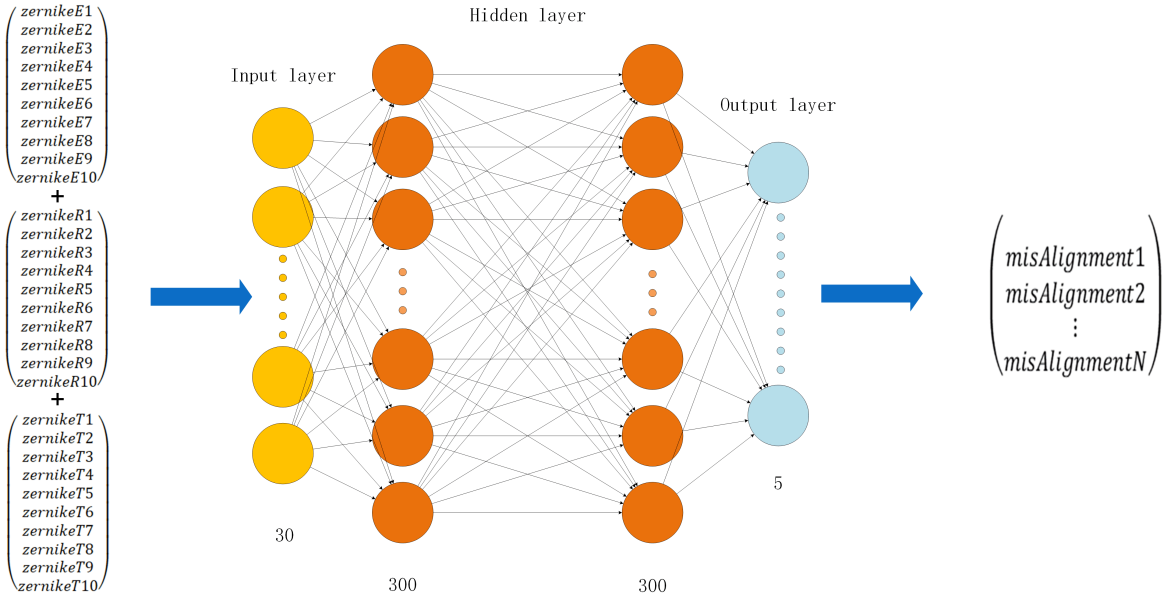


**Fig 9** Establishing a nonlinear neural network model to map the PSF (Zeiss fitting) to telescope misalignment. The network architecture of the coarse neural network and the fine neural network. The architecture of the two networks is identical. The number of nodes for the input layer is 30; each hidden layer contains 300 nodes; the output layer contains eight nodes.

It proposes for the first time a development framework for embodied intelligent telescopes with three stages of state evolution and designs a roadmap for the intelligent evolution of telescopes:

1. Theoretical model-driven stage: In the absence of real data, high-fidelity simulation data is generated through full-process imaging simulation (Figure 2) to provide an initial training set for active optical control. Photon Monte Carlo methods accurately reproduce atmospheric turbulence (von Karman spectrum), optical system reflection/refraction (iterative ray tracing), and detector response (electric field-dependent diffusion model), verifying the interpretability advantage of physical models in complex environments.

2. Multimodal model replacement stage: As real data accumulates, digital twin technology (Figure 3) compares simulation and measurement data to quantify the system deviation of the theoretical model, driving the gradual replacement of the theoretical model by the surrogate model.

3. Data-driven dominant stage: The end-to-end multimodal model directly generates control

commands, and this progressive evolution mechanism provides a scalable path for the autonomous operation and maintenance of large telescopes.

This section introduces breakthroughs in optical performance sensing technology, specifically measuring the physical quantities of the point spread function (PSF) and wavefront aberration data. The PSF model sensing primarily includes:

1. KSB parameterization model. This model first revealed the nonlinear mapping relationship between secondary mirror eccentricity/tilt errors and the ellipticity of the full-field PSF (Figure 4), providing a physical basis for misalignment diagnosis.

2. DAE-PSF reconstruction technology improves the SSIM of reconstructed PSF under low signal-to-noise ratio conditions compared to principal component analysis (PCA) (Figure 6), overcoming the limitations of the PCA method's symmetry assumption and laying the foundation for aberration separation in strongly interfered environments.

Wave aberration sensing primarily involves multi-field curvature wavefront sensing for spatial heterogeneity in field aberrations of fast-focal-ratio telescopes. The four-field curvature wavefront sensor (Figure 7) achieves this through defocus plane intensity differences:

1. Frequency domain filtering (FFT iterative method) achieves high-precision curvature signal reconstruction;

2. Multi-field layout enhances the stability of misalignment solution (sensor $\geq 3$ when the sensitivity matrix is full rank), and the star light utilization rate in the edge field of view is increased by $40\%$.

Embodied intelligent telescopes achieve a closed-loop system of optical performance perception, state inference, and active correction for the first time through the synergy of the interpretability of physical models and data-driven adaptability. A three-stage evolutionary framework ad-

dresses the challenge of transitioning from simulation to real-world applications, with multi-field sensing and the DAE-PSF model significantly enhancing the decoupling capability of image quality degradation. With the development of large-scale models and edge computing, this technology will drive the evolution of telescopes from ”precision instruments” to ”environmentally interactive intelligent agents,” providing core support for unmanned observations in time-domain astronomy.

*References*


1 R. Gilmozzi and J. Spyromilio, “The european extremely large telescope (e-elt),” *The Messenger* **127**(11), 3 (2007).

2 D. Popovic, P. Di Marcantonio, P. Santin, *et al.*, “Eso-vlt/flames: control software for a multi-object observing facility,” in *Advanced Software, Control, and Communication Systems for Astronomy*, **5496**, 107–118, SPIE (2004).

3 M. Ravensbergen, R. Merino, and C. P. Wang, “Encoders for the altitude and azimuth axes of the vlt,” in *Telescope Control Systems*, **2479**, 322–328, SPIE (1995).

4 D.-Q. Su and X.-Q. Cui, “Active optics in lamost,” *Chinese journal of Astronomy and Astrophysics* **4**(1), 1 (2004).

5 R. K. Tyson and B. W. Frazier, *Principles of adaptive optics*, CRC press (2022).

6 B. Ma, Y. Hu, Z. Shang, *et al.*, “Automation of the ast3 optical sky survey from dome a, antarctica,” *Monthly Notices of the Royal Astronomical Society* **496**(3), 2768–2775 (2020).

7 R. Chrisley, “Embodied artificial intelligence,” *Artificial intelligence* **149**(1), 131–150 (2003).

8 J. F. Crenshaw, A. J. Connolly, J. E. Meyers, *et al.*, “Using ai for wave-front estimation with the rubin observatory active optics system,” *The Astronomical Journal* **167**(2), 86 (2024).

9 P. Jia, X. Li, Z. Li, *et al.*, "Point spread function modelling for wide-field small-aperture telescopes with a denoising autoencoder," *Monthly Notices of the Royal Astronomical Society* **493**(1), 651–660 (2020).

10 J. Peterson, J. Jernigan, S. Kahn, *et al.*, "Simulation of astronomical images from optical survey telescopes using a comprehensive photon monte carlo approach," *The Astrophysical Journal Supplement Series* **218**(1), 14 (2015).

11 Z. Wu, Y. Zhang, R. Tang, *et al.*, "Machine learning for improving stellar image-based alignment in wide-field telescopes," *Research in Astronomy and Astrophysics* **22**(1), 015008 (2022).

12 B. Xin, C. Claver, M. Liang, *et al.*, "Curvature wavefront sensing for the large synoptic survey telescope," *Applied Optics* **54**(30), 9045–9054 (2015).

13 Y. Zhang and Y. Qi, "A type of displacement actuator applied on lamost," in *Advanced Software and Control for Astronomy II*, **7019**, 768–773, SPIE (2008).

14 J. Zariski, K. M. Kratter, S. E. Logsdon, *et al.*, "Deep learning solutions to telescope pointing and guiding," in *Software and Cyberinfrastructure for Astronomy VIII*, **13101**, 307–320, SPIE (2024).

15 J. Ngiam, A. Khosla, M. Kim, *et al.*, "Multimodal deep learning.," in *ICML*, **11**, 689–696 (2011).

16 J. Peterson, J. G. Jernigan, A. Dutta, *et al.*, "Phosim: The photon simulator," in *American Astronomical Society Meeting Abstracts*, **237**, 541–13 (2021).

17 J. Peterson, E. Peng, C. Burke, *et al.*, "Deformation of optics for photon monte carlo simulations," *The Astrophysical Journal* **873**(1), 98 (2019).

18 F. Patat, O. Ugolnikov, and O. Postylyakov, "Ubvri twilight sky brightness at eso-paranal," *Astronomy & Astrophysics* **455**(1), 385–393 (2006).

19 A. N. Kolmogorov, "The local structure of turbulence in incompressible viscous fluid for very large reynolds," *Numbers. In Dokl. Akad. Nauk SSSR* **30**, 301 (1941).

20 A. V. Filippenko, "The importance of atmospheric differential refraction in spectrophotometry.," *Publications of the Astronomical Society of the Pacific* **94**(560), 715 (1982).

21 K. Rajkanan, R. Singh, and J. Shewchun, "Absorption coefficient of silicon for solar cell calculations," *Solid-State Electronics* **22**(9), 793–795 (1979).

22 D. M. Botin-Sanabria, A.-S. Mihaita, R. E. Peimbert-Garcia, *et al.*, "Digital twin technology challenges and applications: A comprehensive review," *Remote Sensing* **14**(6), 1335 (2022).

23 N. Kaiser, G. Squires, and T. Broadhurst, "A method for weak lensing observations," *arXiv preprint astro-ph/9411005* (1994).

24 Z. Wu, Y. Zhang, J. Cong, *et al.*, "Multi-field split curvature wavefront sensing and its application in the large field survey telescope," *Research in Astronomy and Astrophysics* **24**(8), 085006 (2024).

25 F. Roddier and C. Roddier, "Wavefront reconstruction using iterative fourier transforms," *Applied Optics* **30**(11), 1325–1327 (1991).

26 M. Liang, V. Krabbendam, C. F. Claver, *et al.*, "Active optics in large synoptic survey telescope," in *Ground-based and Airborne Telescopes IV*, **8444**, 1524–1536, SPIE (2012).

27 D. L. Crawford *et al.*, "The construction of large telescopes," (1966).

28 D. Su, "The researches of the coma of malalignment two-mirror system and infrared chopping secondary mirror system," *Acta Astronomica Sinica* **30**(1), 106 (1989).

29 R. V. Shack and K. Thompson, "Influence of alignment errors of a telescope system on its aberration field," in *Optical Alignment I*, **251**, 146–153, SPIE (1980).

30 H. Lee, G. B. Dalton, I. A. Tosh, *et al.*, "Computer-guided alignment i: Phase and amplitude modulation of alignment-influenced optical wavefront," *Optics express* **15**(6), 3127–3139 (2007).

31 S. Chueca, A. Marín-Franch, A. J. Cenarro, *et al.*, "Curvature wavefront sensing performance simulations for active correction of the javalambre wide-field telescopes," in *Modern Technologies in Space-and Ground-based Telescopes and Instrumentation II*, **8450**, 173–187, SPIE (2012).

32 o. Zhang, "Optical design based on the delano diagram method and study of the antarctic near-infrared survey telescope," (2016).

33 G. Luppino and N. Kaiser, "Detection of weak lensing by a cluster of galaxies at z= 0.83," *The Astrophysical Journal* **475**(1), 20 (1997).

34 Z. Chen, M. Jin, Y. Deng, *et al.*, "Improvement of a deep learning algorithm for total electron content maps: Image completion," *Journal of Geophysical Research: Space Physics* **124**(1), 790–800 (2019).