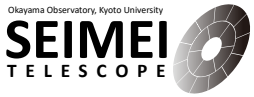
博士学位論文

Development of the point-diffraction interferometer wavefront sensor for extreme adaptive optics

(極限補償光学のための点回折干渉計型波面センサの開発)

Ryo TSUKUI
Department of Astronomy, Kyoto University
津久井 遼
京都大学大学院 理学研究科 物理学・宇宙物理学専攻 宇宙物理学分野
(学籍番号：0560-32-1568)

January 2023
令和 5 年 1 月



Abstract

Direct observation of exoplanets is crucial for solving the following two problems: clarifying the mechanisms of planet formation and searching for extraterrestrial biological activities. Potential targets have small angular separations ($\sim 0.01 - 0.1$ arcsec) from their host stars and small planet-to-star contrasts ($\sim 10^{-7}$). Thus, direct observation requires a high angular resolution and high contrast. Such observation with ground-based telescopes is affected by the Earth's atmospheric turbulence. This is because the turbulence causes wavefront aberrations, which scatter the host star's light to degrade the angular resolution and contrast. Therefore, wavefront correction with adaptive optics (AO) is necessary. AO measures the incident wavefront aberration with a wavefront sensor (WFS) and corrects the aberration with a deformable mirror (DM), controlled by a real-time controller (RTC).

A highly accurate wavefront correction is required to obtain a 10^{-7}-level final contrast. Such correction can be achieved with extreme adaptive optics (ExAO), which has ~ 40000 measurement/correction points in a 30-m telescope aperture and runs at ~ 5 kHz. ExAO requires a high-performance WFS with the following properties:

- high efficiency: a small measurement error with a limited number of photons,
- high-speed capabilities: low calculation cost and a small readout region for wavefront sensing,
- a large dynamic range.

As for high efficiency, ExAO favors phase sensors, such as a fixed pyramid WFS and a Zernike WFS. However, These current phase sensors have room to improve high-speed capabilities and dynamic ranges.

This thesis describes the development of a new phase sensor named b-PDI (birefringent point-diffraction interferometer). The b-PDI utilizes birefringent crystal as its key optical element. The following properties are predicted with simulations and demonstrated with experiments:

- high efficiency comparable to other phase sensors,
- high-speed capabilities with the lowest level of calculation cost and the smallest level of readout region width for wavefront sensing,

- a larger dynamic range than other phase sensors.

Contents

Chapter 1

Introduction: Direct Observation of Exoplanets

Modern astronomy has two challenging problems: (1) clarifying the mechanisms of planet formation and (2) searching for extraterrestrial biological activities. Direct observation of exoplanets can bring essential information to solve these problems. Exoplanets are the planets orbiting stars other than the Sun. After the first detection of the planet around a main-sequence star by Mayor and Queloz in 1995 [1], about 5000 exoplanets [2] have been detected mostly through indirect observation such as radial velocity measurement. Indirect observation reveals the physical properties of exoplanets, such as the mass and the radius, but cannot bring crucial information to solve the problems above. This chapter describes the necessity for direct observation, in which exoplanets and their host stars are spatially resolved with telescopes, and raises the observational characteristics of potential targets.

1.1 Clarifying the mechanisms of planet formation

So far, two planet formation models have been presented: the disk instability (hot-start) model [3] and the core accretion (cold-start) model [4, 5, 6]. Each model predicts different luminosity evolution of gas-giant planets (Fig. 1.1). According to the disk instability model, planets are formed through gravitational collapse via the instability of the protoplanetary disk. This process enables the gas to retain much of its initial heat. Thus, the gas-giant planets just after the formation are hot ($\gtrsim$ 1000 K [7]) and luminous at the infrared band due to a strong thermal emission. On the other hand, according to the core accretion model, gas-giant planets are formed through gas accretion

onto solid cores ($\sim$ 10 Earth masses). The process leads the accreting gas to radiate away much of its initial heat due to an accretion shock [8, 9]. Thus, the gas-giant planets just after the formation are relatively cold ($\lesssim$ 1000 K [7]) and less luminous.

The two models are validated by comparing the observed luminosity of gas-giant exoplanets with the predicted values. Detailed comparison requires the infrared luminosity of gas giants with known ages ($\lesssim$ 0.1 Gyr) and masses. The age is considered the same as the age of its host star. The mass is determined by indirect observations such as astrometry and radial velocity measurement. The infrared luminosity is measured only by direct photometric observation. Therefore, direct observation at the infrared band is essential to clarify the mechanisms of planet formation.

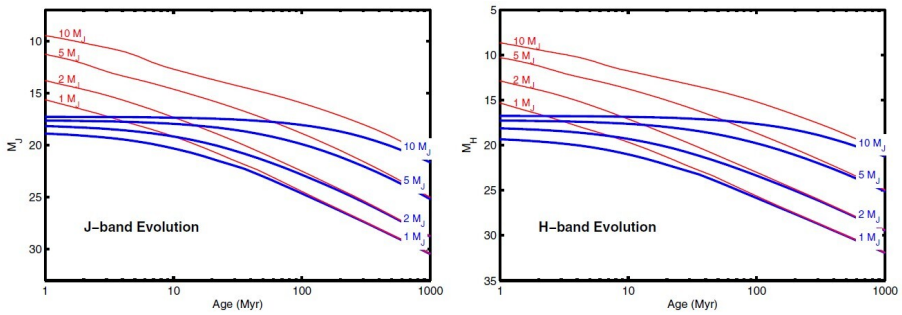


Figure 1.1: The figures from Spiegel & Burrows [7] showing the luminosity evolution of gas-giant planets under two models. Left: evolution of J-band absolute magnitude. Right: evolution of H-band absolute magnitude. Red solid lines: the disk instability (hot-start) model. Blue solid lines: the core accretion (cold-start) model. Gas-giant planets with masses of 1, 2, 5, and 10 Jupiter mass (M_J) are assumed.

1.2 Searching for extraterrestrial biological activities

In recent years, potentially habitable exoplanets have been detected through indirect observations. They are rocky planets and orbiting within the habitable zones around their host stars. Such planets are expected to maintain liquid water and biological activities. Representative examples are the planets around the M-type star TRAPPIST-1 [10].

Detection of biosignatures supports biological activities on exoplanets. Examples of biosignature are atmospheric O_2 and CH_4 [11, 12], which can

be produced by biological activities. In particular, simultaneous detection of O_2 and CH_4 provides a strong support for the existence of life [12].

These substances produce absorption lines in the exoplanet's reflected light and emitted light. The wavelengths of typical O_2 and CH_4 absorption lines are respectively 1.27 μm and $\sim$ 1.6 μm (Fig. 1.2), around which reflected light is prominent. Thus, the absorption lines could be detected by direct spectroscopic observation of the reflected light at the near-infrared bands. The direct spectroscopy in this band is also useful for disctiminating the false positives, as shown in Fig. 1.2.

Note that the atmospheric composition of exoplanets can also be searched by spectroscopy of the transmitted light. However, the transmission spectroscopy suffers from low signal-to-noise (S/N) ratios because the strong light arriving directly from the star is mixed into the transmitted light and becomes noise. In contrast, direct spectroscopic observations can yield relatively high S/N ratios because the light from the planet and the star is physically isolated [13]. In addition, direct observations of reflected light can access deeper atmosphere of planets than transmission spectroscopy [14].

1.3 Characteristics of potential targets

Exoplanets are located at small angular separation ($\ll 1$ arcsec) from their host star. In addition, its luminosity is several orders of magnitude less than that of its host star. In other words, the planet/star luminosity ratio (contrast) is small. This property is referred to as "high contrast" This section specifies the angular separations and contrasts of the potential targets.

1.3.1 Gas-giant planets

In order to verify the models of planet formation, the targets should be young gas-giant planets with known masses. Matsuo et al. focused on the gas-giant planets [17] whose masses are determined by GAIA astrometry mission. Astrometry is an effective method to determine the masses of young planets. This is because the method is less subject to the activity of young host stars than radial velocity measurement. Matsuo et al. estimated the luminosity and contrast of the planets within the discovery space of GAIA. Figure 1.3 shows the result. The assumed host stars are solar-type with masses of $1M_\odot$ [1]. With a $\sim$ 20-μarcsec astrometry accuracy and a

[1] $M_\odot$: Solar mass.

5-year mission lifetime, GAIA can detect planets with masses of $\gtrsim 0.1\ M_J$ [2] and orbital periods of less than several years [18]. The periods correspond to angular separations of $\lesssim 0.6$ arcsec from the host stars at 5 pc from the Earth. The contrast of such planets are $\gtrsim 10^{-10}$. In particular, the planets at $> 5\ M_J$ have contrasts of $10^{-6} - 10^{-7}$.

1.3.2 Potentially habitable planets

The targets are earth-sized planets orbiting within habitable zones. Guyon et al. estimated the angular separations and contrasts of hypothetical Earth analogs at the habitable zones of the stars within 30 pc from the Earth [19]. Figure 1.4 shows the result. The estimated angular separations and contrasts are grouped by the spectral types of the host stars. In the case of G-type host stars, the angular separations are relatively large (~ 0.1 arcsec) because the habitable zones are distant from hot host stars. The contrasts are $\sim 10^{-10}$ because of the high luminosity of the host stars. In the case of M-type host stars with low temperature, in contrast, the angular separations are small (~ 0.01 arcsec), and the contrasts are relatively moderate ($\sim 10^{-7}$).

1.4 Summary of this chapter

Direct observation of exoplanets in near-infrared wavelengths brings crucial information about planet formation and exraterrestrial biologiacl activities. Potential targets have small angular separations ($\sim 0.01 - 0.1$ arcsec) from their host stars and small planet/star luminocity ratios ($\sim 10^{-7} - 10^{-10}$). The latter property is referred to as "high contrast."

[2] M_J: Jupiter mass.

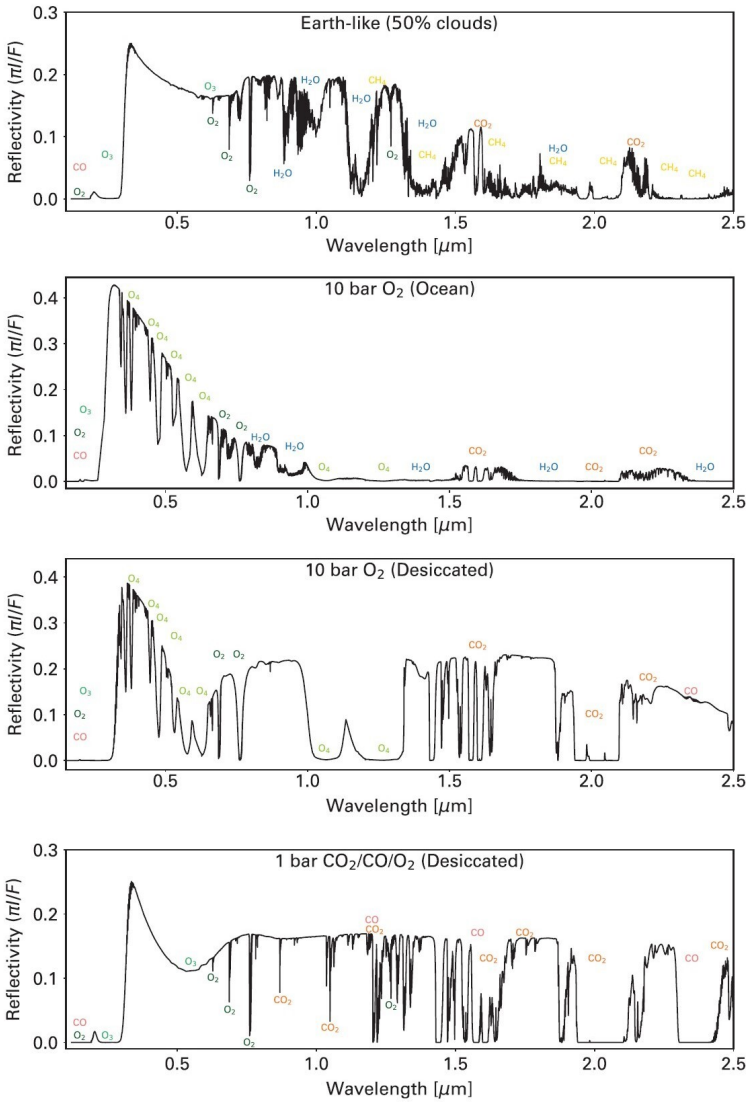


Figure 1.2: The figures from Meadows et al. [15] showing the predicted reflectivity spectra of the potential planet Proxima Centauri b. The topmost figure assumes Earth-like atmosphere containing biological O_2 and CH_4. The second to fourth figures from the top are examples of false positives with abiotic O_2. The second and third assume the atmosphere with massive O_2 produced via the ocean loss and the succeeding photolysis of H_2O. These cases show broad O_4 absorption lines induced by O_2-O_2 collisions [16]. The fourth assume the atmosphere with O_2 produced via the photolysis of CO_2. This case shows prominent CO lines [16]. These features are used to discriminate false positives.

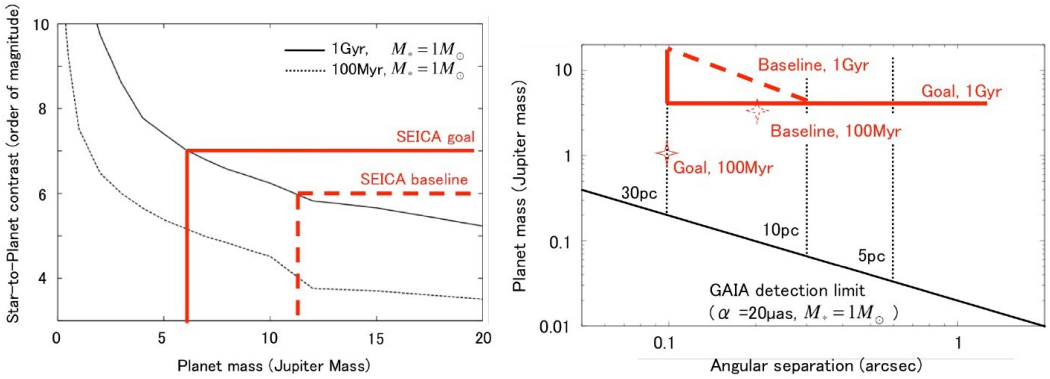


Figure 1.3: The figures cited from Matsuo et al. [17] Left: the estimated star-to-planet contrast plotted against the mass of gas-giant planets. The black solid and dashed lines assume the planetary age of 1 Gyr and 100 Myr, respectively. Right: Discovery space of GAIA. The horizontal axis is the angular separation between the planet and its host star. The vertical axis is the planetary mass normalized by Jupiter's mass. The black solid line represents the lower limit of the discovery space, assuming GAIA's astrometry accuracy of 20 μarcsec, mission lifetime of 5 years, and the host star's mass of $1M_\odot$. The black dotted lines are the outer limits on the detectable angular separation for various distances. The red lines in both figures are the potential discovery space of the high-contrast instrument SEICA described later.

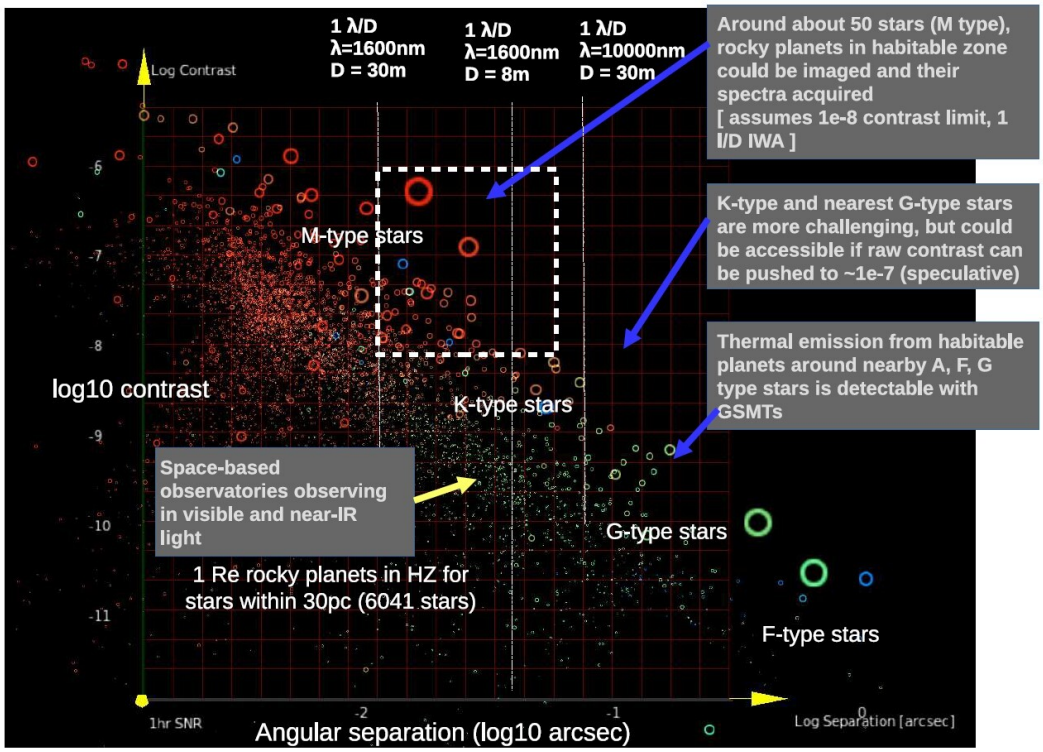


Figure 1.4: The figure from Guyon et al. [19] showing the samples of potentially habitable planets. The horizontal axis is the angular separation of the planet and its host star. The vertical axis is the logarithmic contrast.

Chapter 2

Atmospheric Turbulence and Adaptive Optics

The previous chapter showed that the direct observation of exoplanets requires high angular resolution ($\sim 0.01 - 0.1$ arcsec) and high contrast ($\sim 10^{-7} - 10^{-10}$). However, the turbulence in the Earth's atmosphere has negative impacts on high-resolution and high-contrast observation with ground-based telescopes. This chapter characterizes the atmospheric turbulence and introduces the relevant parameters. In addition, this chapter overviews adaptive optics (AO), which is a technology to overcome the turbulence.

2.1 Atmospheric turbulence and wavefront aberrations

The Earth's atmosphere is always turbulent. This turbulence forms fluctuations in the refractive index distribution in the atmosphere and disturbs the light from stars. As a result, the telescopic image of the host star is disturbed, degrading the angular resolution and contrast around it. This section outlines this process based on Hardy [20] and Guyon [13].

2.1.1 Description of atmospheric turbulence

The shear of the wind velocity in the Earth's atmosphere causes turbulence. Kolmogorov presented the model describing the velocity distribution of turbulent fluid [21]. In the model, turbulence is formed as large-scale (outer scale L_0) disturbances, which gradually break down into smaller structures.

When the structure size reaches the minimum scale (inner scale l_0), the kinetic energy of the disturbance dissipates as heat. Kolmogorov found that isotropic and homogeneous fluid in turbulence has the velocity distribution $v(x)$, which satisfies the following structure function $D_v(r)$:

$$D_v(r) = \langle [v(x) - v(x+r)]^2 \rangle = C_v^2 r^{2/3}. \tag{2.1}$$

In the function, r is the distance between two arbitrary points in the fluid and is restricted as $l_0 < r < L_0$. In the Earth's atmosphere, $L_0 = 1 - 100$ m and $l_0 = 1 - 10$ mm [22]. The operator $\langle\rangle$ means an ensemble average with respect to the position x. The constant C_v denotes the strength of the turbulence.

The turbulence mixes air masses with different altitudes from the Earth's surface. The mixing causes the fluctuations in temperature because the air masses with different altitudes have different temperatures. Thus, the fluctuations are prominent under the altitude of ~ 10 km, where the temperature gradient is large [20]. The turbulence-induced fluctuations in temperature have a similar scale to the turbulence in the background [23]. Thus, the temperature distribution $T(x)$ satisfies the following structure function $D_T(r)$:

$$D_T(r) = \langle [T(x) - T(x+r)]^2 \rangle = C_T^2 r^{2/3}, \tag{2.2}$$

where C_T is constant.

The temperature fluctuations produce the fluctuations in the refractive index. The refractive index n of the air depends on its temperature T and pressure P [20] as follows:

$$N = n - 1 = (77.6 \times 10^{-6}) P/T. \tag{2.3}$$

The local fluctuations in the pressure P rapidly diffuse at the speed of sound and are smoothed out. Hence, they are negligible when compared with the fluctuations in the temperature T. Therefore, the refractive index fluctuations ΔN are mainly caused by the temperature fluctuations ΔT:

$$\Delta N = (-77.6 \times 10^{-6}) \frac{P}{T^2} \Delta T. \tag{2.4}$$

Under the relationship, the refractive index distribution $N(x)$ follows the

structure function $D_N(r)$:

$$D_N(r) = \langle [N(x) - N(x+r)]^2 \rangle \tag{2.5}$$

$$= \left(\frac{\Delta N}{\Delta T}\right)^2 D_T(r) \tag{2.6}$$

$$= \left(\frac{\Delta N}{\Delta T}\right)^2 C_T^2 r^{2/3} \tag{2.7}$$

$$= C_N^2 r^{2/3}. \tag{2.8}$$

The term C_N^2 in the function is described as follows:

$$C_N^2 = \left(\frac{\Delta N}{\Delta T}\right)^2 C_T^2. \tag{2.9}$$

Generally, C_N^2 is a function of the altitude H and denotes the strength of the fluctuation at the altitude.

2.1.2 Atmospheric wavefront aberration

The refractive index fluctuations disturb the wavefront of the light (wavelength: λ) passing through the atmosphere. In the absence of the fluctuations, light from sufficiently distant point-sources is considered as a plane wave. The refractive index fluctuations deform the plane wave and produce wavefront aberration, which is defined as deviation from the plane wave. The atmospheric wavefront aberration $W(x)$ is formulated by integrating Eq. 2.8 along the optical path from the sky to the ground [22]. The integration results in the following structure function $D_\phi(r)$ of the wavefront phase $\phi(x)$ [rad] ($= 2\pi W(x)/\lambda$):

$$D_\phi(r) = \langle [\phi(x) - \phi(x+r)]^2 \rangle \tag{2.10}$$

$$= 2.914 k^2 (\sec z) r^{5/3} \int_h C_N(h)^2 dh \tag{2.11}$$

$$= 6.88 \left(\frac{r}{r_0}\right)^{5/3}, \tag{2.12}$$

where z is a zenith angle of the optical path, and $k = 2\pi/\lambda$. The term r_0 is the Fried parameter [24], which is defined as:

$$r_0 = \left[0.423 k^2 (\sec z) \int_h C_N(h)^2 dh\right]^{-3/5}. \tag{2.13}$$

Using the Fried parameter r_0, the phase variance σ_ϕ^2 [rad^2] over the telescope aperture of diameter D is expressed by the following equation [25]:

$$\sigma_\phi^2 = 1.0299 \left(\frac{D}{r_0}\right)^{5/3}. \tag{2.14}$$

Therefore, r_0 is interpreted as the aperture size over which the phase variance is ~ 1 rad^2. In the case with $D = 30$ m and $r_0 = 0.15$ m, $\sigma_\phi = 84$ rad, which corresponds to 6.7 μm at $\lambda = 0.5$ μm.

The Fried parameter r_0 is used to define the parameters related to the time and angular dependence of the wavefront aberration. First, coherence time t_0 is a parameter that characterizes the time dependence of atmospheric wavefront aberration. The atmospheric turbulence is swept laterally by the average wind speed V_0 and crosses over the line of sight faster than the turbulence itself develops. Due to this property, the time evolution of atmospheric turbulence is approximated by the frozen flow assumption [26]; a screen with fixed turbulence crosses the line of sight with the velocity V_0. When the turbulence is localized in one layer, the wavefront phase change $\sigma_\phi(\Delta t)$ [rad RMS [1]] with time Δt is described by the following equation [13]:

$$\sigma_\phi(\Delta t)^2 = \langle [\phi(x,t) - \phi(x,t+\Delta t)]^2 \rangle = \left(\frac{\Delta t}{t_0}\right)^{5/3}, \tag{2.15}$$

where the coherence time t_0 is defined as:

$$t_0 = 0.31 \frac{r_0}{V_0}. \tag{2.16}$$

Thus, t_0 stands for the average time interval in which the change in the wavefront aberration reaches 1 rad RMS over the telescope aperture. When $r_0 = 0.15$ m and $V_0 = 10$ m/s, $t_0 = 4.7$ ms.

Second, the isoplanatic angle θ_0 is a parameter that characterizes the angular dependence of the wavefront aberration. When the turbulence is localized in one layer, the wavefront phase change $\sigma_\phi(\theta)$ [rad RMS] due to the difference in line-of-sight direction $\Delta\theta$ is described by the following equation [20]:

$$\sigma_\phi(\theta)^2 = \langle [\phi(x,\theta) - \phi(x,\theta+\Delta\theta)]^2 \rangle = \left(\frac{\Delta\theta}{\theta_0}\right)^{5/3}, \tag{2.17}$$

[1] Root-mean-square value over the telescope aperture

where the isoplanatic angle θ_0 is defined as:

$$\theta_0 = 0.31(\cos z)\frac{r_0}{H}. \tag{2.18}$$

Thus, θ_0 stands for the difference in line-of-sight direction with which the change in the wavefront aberration reaches 1 rad RMS over the telescope aperture. When $r_0 = 0.15$ m and $H = 10$ km, $\theta_0 \simeq 1$ arcsec at the zenith ($z = 0$).

2.1.3 Effect on the image quality

Wavefront aberration disturbs the telescope's point-spread function (PSF), which is the image of a point source such as a star. With an aberration-free telescope, plane wave yields a diffraction-limited image (Fig. 2.1(a)), whose angular core size is $\sim \lambda/D$ [rad]. When $\lambda = 1.2$ μm and $D = 30$ m, $\lambda/D \sim 0.008$ arcsec. Hence, the angular resolution is ~ 0.008 arcsec. In contrast, the atmospheric wavefront aberration yields an aberrated PSF called a seeing-limited PSF (Fig. 2.1(c)), whose angular size is $\sim \lambda/r_0$ [rad]. When $\lambda = 1.2$ μm and $r_0 = 0.15$ m, $\lambda/r_0 \sim 1$ arcsec. In the direct observation of exoplanets, the atmospheric wavefront aberration degrades the angular resolution and contrast by scattering the host star's light into the areas where planets exist.

An aberrated PSF is interpreted as a collection of speckles, which are point images with diffraction-limited image sizes. There is a relationship between the position of the speckle and the pattern of wavefront aberration. This is explained by Fraunhofer diffraction. That is, the electric field of the PSF is the Fourier transform of the electric field of the observed light with wavefront aberration. Thus, an wavefront aberration with a single spatial frequency create a pair of speckles (Figure2.2). The higher frequency components create speckles at a greater distance from the PSF center. Wavefront aberrations with various spatial frequencies create an aberrated PSF by creating numerous speckles. This effect has been quantified by Guyon [27]. The wavefront phase $\phi(\mathbf{u})$ ($\mathbf{u}$: position (x, y)) at the telescope aperture is expressed as a sum of sinusoidal components:

$$\phi(\mathbf{u}) = \sum_i \frac{2\pi h_i}{\lambda} \cos(2\pi \mathbf{f}_i \mathbf{u} + \psi_i), \tag{2.19}$$

where each component has an amplitude h_i, a spatial frequency $\mathbf{f}_i$, and a translation ψ_i. When the amplitude $h_i \ll \lambda$, the resulting PSF is approxi-

mated by the following equation:

$$I(\mathbf{a}) \simeq I_0(\mathbf{a}) + \sum_i \left(\frac{\pi h_i}{\lambda}\right)^2 [I_0(\mathbf{a} + \mathbf{f}_i\lambda) + I_0(\mathbf{a} - \mathbf{f}_i\lambda)], \tag{2.20}$$

where $\mathbf{a}$ [rad] is the angular coordinate on the celestial sphere, and I_0 is the diffraction-limited PSF. The terms after the $\sum_i$ stand for a pair of speckles corresponding to a sinusoidal component. This equation relates the amplitude of the wavefront aberration and the intensity of the speckles.

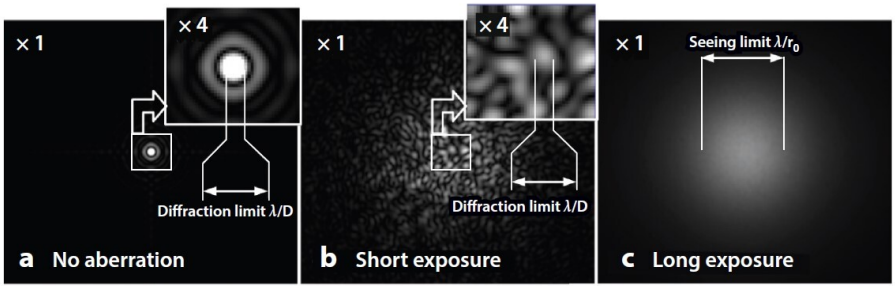


Figure 2.1: Simulated PSFs with an 8-m telescope at the infrared wavelength, cited from Guyon. [13] (a) a diffraction-limited PSF without wavefront aberration. Its angular core size is $\sim \lambda/D$. (b) a short-exposure image of a seeing-limited PSF, consisting of numerous speckles. The speckles have diffraction-limited size $\sim \lambda/D$. (c) a long-exposure image of a seeing-limited PSF, whose angular size is $\sim \lambda/r_0$.

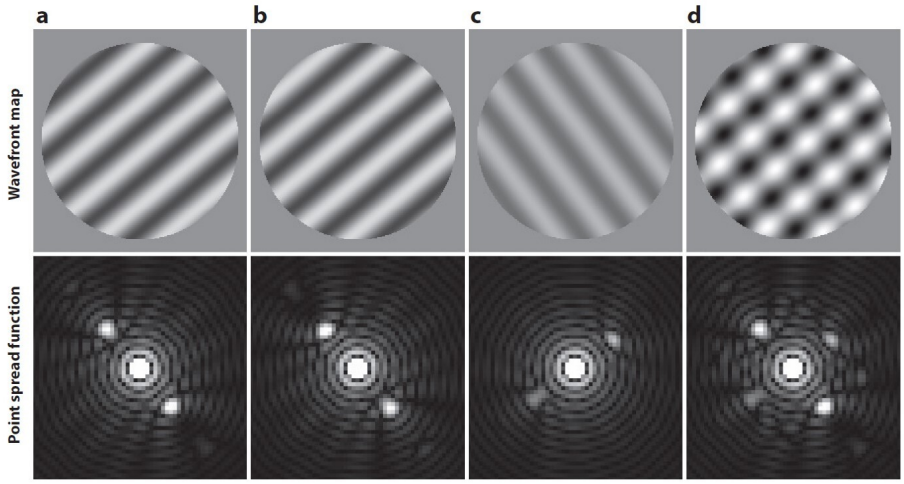


Figure 2.2: The relationship between the wavefront aberration (top row) and the speckle position (bottom row), cited from Guyon. [13]

2.2 Adaptive optics

The adaptive optics (AO) system corrects the atmospheric wavefront aberration to improve the angular resolution and contrast. This section outlines the components and error sources of the AO.

2.2.1 Overview

A general AO consists of a wavefront sensor (WFS), a real-time controller (RTC), and a deformable mirror (DM). Figure 2.3 shows a closed-loop AO configuration. In the configuration, the WFS converts the incident wavefront aberration into light intensity distributions (signals), which are read out with detectors. The RTC calculates the wavefront aberration based on the WFS signal and drives the DM to cancel the aberration. This sequence of operations (AO loop) is repeated to correct time-varying wavefront aberrations continuously. While the correction continues, the WFS measures the residual aberration of the corrected wavefront, and the DM is driven to reduce the residual aberration further. A general AO measures the wavefront aberration with visible light and provides corrected infrared light to downstream science instruments such as imaging cameras or spectrographs.

The WFS has a finite number of measurement points (subapertures), which sample the incident wavefront aberrations. The principles of typical WFSs are specified in Chapter 4. On the other hand, the DM has a

thin-film mirror surface that is driven by actuators on the back side, which are typically composed of piezoelectric elements or micro-electro-mechanical systems (MEMS).

Typical parameters for AO performance are spatial resolution N_{SA} of the wavefront and loop speed F_{AO}. N_{SA} is the number of the WFS subapertures or the number of the DM actuators. N_{SA} is related to the fitting error described below and is determined based on r_0. F_{AO} is the number of AO loops per second. F_{AO} is related to the temporal error and WFS noise error described below and is determined with reference to t_0.

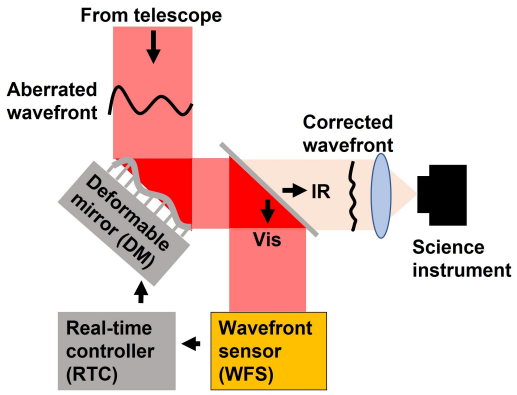


Figure 2.3: Schematic drawing of a closed-loop adaptive optics. Vis: visible light. IR: infrared light.

2.2.2 Errors in adaptive optics

The errorrs listed below result in a residual aberration of the AO-corrected wavefront. The residual aberration is a root-sum-square of these errors because they are independent of each other. The residual aberration forms speckles, as shown by Eq. 2.20, and limits the angular resolution and contrast. Therefore, the error sources must be considered in designing the AO with a designated angular separation and contrast.

- **Fitting error**: According to Kolmogorov's model, the spatial scale of wavefront aberration spans from L_0 ($\gtrsim 1$ m) to l_0 ($\gtrsim 1$ mm). However, correcting the wavefront aberration down to the smallest scale is not practical. One reason for this is the limited number of DM actuators that can be manufactured. Therefore, high-frequency wavefront aberrations that exceed the DM's Nyquist frequency remain

uncorrected. The fitting error σ_{fit} is described by using the actuator spacing d (projected onto the telescope aperture) as follows [28]:

$$\sigma_{fit}^2 = a_{fit}(d/r_0)^{5/3}, \tag{2.21}$$

where the coefficient a_{fit} is dependent on the types of DM surface. The 1-D spatial resolution $N_{1d} = D/d$ is determined with reference to the required value of σ_{fit}.

- **Temporal error**: There is an unavoidable time lag Δt between the WFS measurement and the DM correction. The time lag includes the exposure time of the WFS, the readout time of the detector, and the calculation time of the RTC. The time lag produces a temporal error represented by the expression 2.15. Increasing F_{AO} reduces this error.

- **WFS noise error**: The signals captured by the WFS detectors contain random noise. Typical noise is photon noise and readout noise. Even an ideal detector without readout noise is affected by photon noise. These noises produce errors in the measurement of wavefront aberration. This is the WFS noise error. This error can be reduced by improving the S/N ratio. One specific method is increasing the WFS exposure time by decreasing F_{AO}, which increases temporal error. Therefore, the tradeoff between the temporal and WFS noise errors must be considered in determining the optimal F_{AO}.

- **Anisoplanatic error**: When the WFS's reference star (guide star) and the science target have a separation angle, the turbulence detected by the WFS and the turbulence in the optical path from the science target are different. Therefore, the separation angle $\Delta\theta$ produces the anisoplanatic error represented by the equation 2.17.

- **Chromatic wavefront due to reflactive index**: The refractive index of the atmosphere depends on wavelength, although this is not considered in Section 2.1.2. The refractive index $n(\lambda)$ of dry air at standard conditions is given by [29]:

$$N(\lambda) = (n(\lambda) - 1) \times 10^6 \tag{2.22}$$

$$= 8.34213 \times 10^{-5} + \frac{2.40603 \times 10^{-2}}{130 - \lambda^{-2}} + \frac{1.5997 \times 10^{-4}}{38.9 - \lambda^{-2}}, \tag{2.23}$$

where λ is expressed in μm. Based on this equation, the refractive index of the atmosphere in the visible region differs from that in the

infrared region by $\sim 10^{-5}$. The difference changes the scale of the wavefront aberration in the optical path direction between the visible WFS and the infrared science instrument. This difference in scale creates errors in the aberration correction in the IR region.

- **Chromatic wavefront due to atmospheric dispersion**: The density gradient of the Earth's atmosphere refracts light incident at an oblique angle. The degree of refraction depends on the wavelength. This is atmospheric dispersion. Therefore, visible light used by the WFS and infrared light used by the science instruments take different paths in the atmosphere to reach the telescope. Due to the lateral shift of these paths ($\sim$ 1 cm [13]), the turbulence detected by the WFS does not exactly match the turbulence affecting the science target. This creates errors in the aberration correction.

2.3 Summary of this chapter

The Earth's atmospheric turbulence causes wavefront aberration, which is characterized by the Fried parameter r_0, coherence time t_0, and isoplanatic angle θ_0. The Wavefront aberration distorts the telescope's images, degrading the angular resolution and contrast. AO corrects the wavefront aberration and improves the angular resolution and contrast. The AO design must consider the parameters to reduce the residual aberrations and achieve the designated angular resolution and contrast.

Chapter 3

Extreme Adaptive Optics

To achieve high angular resolution ($0.01 - 0.1$ arcsec) and high contrast ($\sim 10^{-7} - 10^{-10}$), AO needs to correct the wavefront aberrations with high precision. An AO specialized for this purpose is extreme adaptive optics (ExAO). This chapter provides an overview of ExAO and describes the requirements for ExAO and WFS based on the study of an actual direct observation project.

3.1 Overview of extreme adaptive optics

Improving the angular resolution and contrast with AO is equivalent to bringing the aberrated PSF closer to the diffraction-limited PSF. The difference between the aberrated and diffraction-limited PSF is described by the Strehl ratio (SR). The SR is the ratio between the peak intensity of the aberrated PSF and that of the diffraction-limited PSF. Using the RMS value σ_W [μm] of the wavefront aberration, the SR is approximated by the following equation (the extended Marechal approximation [20]):

$$\mathrm{SR} = \exp\{-(2\pi\sigma_W/\lambda)^2\}. \tag{3.1}$$

For example, SR > 0.9 requires $\sigma_W \lesssim \lambda/20$, which corresponds to $\sigma_W \lesssim$ 60 nm at $\lambda = 1.2$ μm. Such highly accurate correction requires a strong reduction of each error in the AO (Section 2.2.2); ExAO is specialized for this purpose and has the following characteristics.

3.1.1 High spatial resolution

ExAO has a high spatial resolution (N_{SA}) of wavefronts to reduce the fitting error. The 1-D spatial resolution is generally set by:

$$N_{1d} \simeq D/r_0. \tag{3.2}$$

When $D = 30$ m and $r_0 = 0.15$ m, $N_{1d} \simeq 200$. Thus, $N_{SA} \sim N_{1d}^2 \sim 40000$.

The DM requires N_{SA} ($\sim N_{1d}^2$) actuators and $\sim$ 10-μm stroke to deal with the incident wavefront aberration expressed by Eq. 2.14. Currently, available DMs are not able to satisfy both two requirements. Thus, a typical ExAO consists of two AO systems connected in series. The first stage ("Woofer") has a small spatial resolution and a large stroke; it corrects low-order and large-amplitude wavefront aberrations. The second stage ("Tweeter") has a large spatial resolution (Eq. 3.2) and a small stroke; it corrects high-order and small-amplitude wavefront aberrations along with low-order residuals of the Woofer.

3.1.2 High speed

ExAO (Tweeter) runs a fast AO loop to reduce the temporal error. The speed F_{AO} is generally set by:

$$F_{AO} \sim 10/t_0, \tag{3.3}$$

which means 10 AO loops per t_0. When $r_0 = 0.15$ m and $V = 10$ m/s, $F_{AO} \sim 2$ kHz.

A high-speed AO loop has the problem of increasing WFS noise error. The error can be suppressed by a WFS with a small measurement error induced by noise. This property is referred to as "high efficiency" and is one of the necessary properties of the Tweeter WFS in ExAO.

3.1.3 Narrow field of view

The host star is used as a guide star in direct observations of exoplanets. The science target planet is located within $\sim$ 0.1 arcsec of the guide star. This separation angle is smaller than the typical isoplanatic angle θ_0. Thus, the anisoplanatic error is negligible.

3.1.4 Supression of the chromatic errors

The chromatic errors stand for the chromatic wavefront due to atmospheric dispersion and refractive index introduced in Section 2.2.2. There are several proposals to attempt the error. The first example is an additional

infrared WFS [19] along with a visible Tweeter WFS. The second example is a Tweeter WFS operated in the infrared [17].

3.2 High-contrast instruments

The ExAO-corrected PSF of the host star has two factors limiting the contrast. The first is the diffraction patterns due to the telescope aperture. The second is the speckles induced by the residual wavefront aberrations of ExAO. These limiting factors are suppressed by coronagraphs and post-process connected downstream of ExAO. The system consisting of these instruments and ExAO is called a high-contrast instrument (HCI) and ultimately achieves a contrast at the $< 10^{-6}$ level.

Figure 3.1 shows the process of the contrast improvement with HCI. (A) Without ExAO-correction, the PSF of the host star is seeing-limited and has a spread of $\sim \lambda/r_0$ (~ 1 arcsec). (B) After ExAO-correction, the PSF core of the host star is converged to the size of $\sim \lambda/D$, improving the angular resolution and contrast. However, the PSF has diffraction patterns and residual speckles. (C) The diffraction patterns are canceled by coronagraphs [30]. The residual speckles are reduced by post-process (not shown in the figure). The post-process includes differential imaging (i.e. [31, 32]) and speckle nulling [33].

3.3 Requirements for high-contrast instruments and extreme adaptive optics

To achieve the target contrast, one must allocate the contrast improvement to each HCI component, as shown in Fig. 3.1. This section specifies the allocated contrast improvement of ExAO based on the two studies of HCI projects.

3.3.1 SEICA instrument

SEICA (Second-generation Exoplanet Imager with Coronagraphic Adaptive optics; Matsuo et al. [17]) is an HCI to be mounted on the Seimei telescope ($D = 3.8$ m) [34]. To observe gas-giant planets, SEICA aims to achieve a contrast of $10^{-6} - 10^{-7}$ at $0.1 - 0.2$-arcsec ($1.2 - 2.4\lambda/D$) angular separation and H-band. Its ExAO is required to achieve a $10^{-3} - 10^{-4}$ contrast.

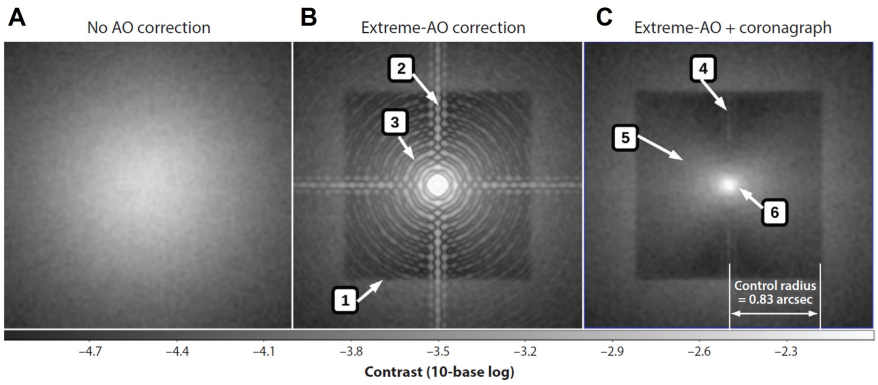


Figure 3.1: The process of the contrast improvement with a simulated HCI, cited from Guyon. [13] The indices (A-C) have been modified from the original figure. (A) The seeing-limited PSF of the host star with a spread of $\sim \lambda/r_0$ (~ 1 arcsec). (B) The ExAO-corrected PSF of the host star. 1: ExAO control radius, which is the angular separation $(N_{1d}/2) \times (\lambda/D)$ corresponding to the Nyquist frequency of the DM. 2: The diffraction pattern due to the spiders (the support structure of the secondary mirror) at the telescope aperture. 3: The diffraction pattern due to the edge of the telescope apertre. (C) The PSF after ExAO and coronagraph. 4, 5, and 6: The collection of the speckles due to the residual aberrations

3.3.2 PSI instrument

PSI (Planetary Systems Imager; Guyon et al., [19, 35]) is an HCI to be mounted on the TMT (Thirty Meter Telescope; $D = 30$ m). To observe potentially habitable planets around M-type stars, PSI aims to achieve a contrast of $\sim 10^{-7}$ at 0.015-arcsec ($1.3\lambda/D$) angular separation and J-band. Its ExAO is required to achieve a $\sim 10^{-5}$ contrast.

3.4 Requirements for wavefront sensors

ExAO requires a high-performance Tweeter WFS to achieve the contrast specified above. As shown in Section 2.2.2, there are two ExAO errors related to the Tweeter WFS: the WFS noise error and the temporal error. These errors create speckles following Eq. 2.20 and limits the contrast. To suppress the contrast limits below the target value under the trade-off nature, the Tweeter WFS is required to have the following properties.

- **High efficiency**: High efficiency stands for a small measurement er-

ror (statistical error) with a certain number of photons. This property suppresses the WFS noise error with a short exposure time, relaxing the trade-off. SEICA [17] and PSI [19, 35] require an efficiency comparable to or better than that of a fixed pyramid WFS described below. Note that this property should be achieved when the wavefront aberration incident to the WFS is around 0. This is because the WFS in a closed-loop AO measures small residual aberrations of DM-corrected wavefronts during most of its operation.

- **High-speed capabilities**: The trade-off studies result in requiring sampling rates ($= F_{AO}$) of $5 - 8.5$ kHz (SEICA [17]) and $1 - 5$ kHz (PSI [19]). These high sampling rates require the following properties.
  - **Low calculation cost**: The RTC calculates the wavefront aberration from the WFS signals. The calculation deals with N_{SA} ($\sim$ 40000) subapertures at ~ 5 kHz. Thus, a low calculation cost is crucial to achieve the high sampling rate.
  - **Small readout region**: The WFS detectors read out the signals with $\propto N_{SA}$ pixels. The frame rates of the detectors must match the required sampling rates. For this purpose, CMOS detectors are potential candidates because they do not need time for a charge transfer, whereas CCD detectors need the transfer. In particular, many scientific CMOS (sCMOS) detectors have a circuit that simultaneously reads out the pixel values in each horizontal line (Fig. 3.2). Hence, the frame rate of sCMOS detectors is further enhanced by minimizing the vertical size of the readout region.
- **High spatial resolution**: Equation 3.2 determines the approximate number of the Tweeter WFS's subapertures. Assuming $r_0 = 0.15$ m, the 1-D spatial resolution $N_{1d} = 25$ and 200 for apertures with $D = 3.8$ m and 30 m, respectively.
- **Large dynamic range**: A Tweeter WFS with a large dynamic range relaxes requirements for an upstream Woofer and stabilizes the operation of ExAO.

As evidence for the required High efficiency and High-speed capability above, the contrast estimates in SEICA [17] are shown in Figure 3.3. The WFS noise error and temporal error limit the contrast to $\sim 10^{-5}$, which satisfies SEICA's requirement for ExAO ($10^{-3} - 10^{-4}$).

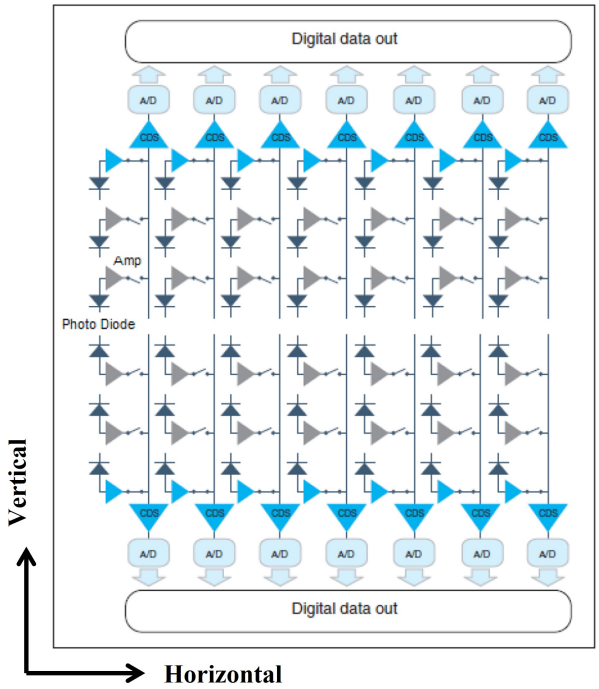


Figure 3.2: Structure of an sCMOS image sensor, cited from the manual of Hamamatsu ORCA-Flash 4.0V2 sCMOS camera. [36] The indicators of the horizontal and vertical directions are added. Each pixel have a combination of a photo diode (the black triangle) and an amp (the right-pointing gray/blue triangle). The circuit simultaneously reads out the pixel values in each horizontal line (row). CDS: correlated double sampling circuit. A/D: analog-to-digital converter.

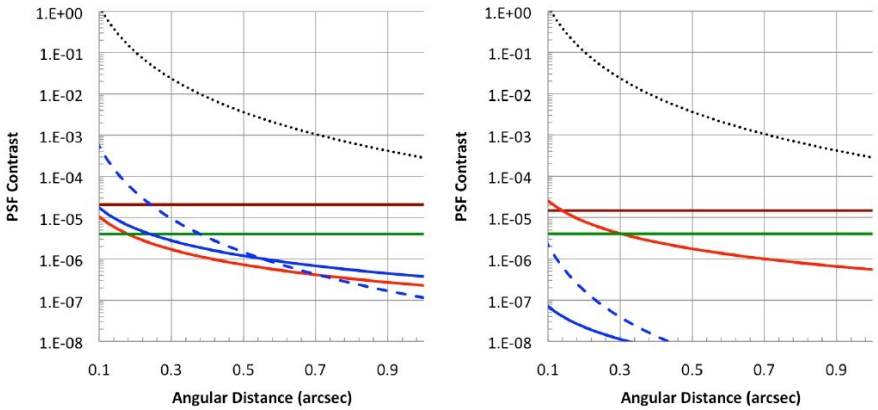


Figure 3.3: The contrast estimates in SEICA, cited from Matsuo et al. [17] The contrast is calculated under the following assumptions: imaging wavelength $\lambda = 1.65$ μm, $D = 4$ m, $r_0 = 10$ cm, $V_0 = 10$ m/s, $N_{1d} = 24$, and a 6-magnitude guide star at the Tweeter WFS's wavelength band. Left: the baseline design in which the Tweeter WFS operates at $F_{AO} = 8.5$ kHz in the visible wavelength band (0.8 ± 0.1 μm). Right: the goal design in which the Tweeter WFS operates at $F_{AO} = 5$ kHz in the infrared wavelength band (1.65 ± 0.15 μm). The red curves: the temporal error. The brown curves: the WFS noise error. The green curves: the fittig error. The blue solid curves: the wavefront chromaticity due to atmospheric dispersion. The blue dashed curves: the wavefront chromaticity due to refractive index. The black dotted curves: without ExAO correction. Note that these figures do not consider the diffracton patterns.

3.5 Summary of this chapter

A high-precision wavefrot correction with ExAO is necessary in direct observations of exoplanets, which have small angular separations ($0.01 - 0.1$ arcsec) and high contrasts ($\sim 10^{-7}$).

The high-contrast instrument (HCI) consists of ExAO, coronagraph, and post-process. The combination of these elements achieves high contrast at a small angular separation. To obtain a final contrast of $\sim 10^{-7}$ at near-infrared wavelengths, ExAO must achieve a contrast of $\sim 10^{-5}$. Therefore, ExAO requires a Tweeter WFS with the following characteristics.

- High efficiency
- High-speed capabilities
  - Low calculation cost
  - Small readout region
- High spatial resolution
- Large dynamic range

Chapter 4

Review of Wavefront Sensors

The previous chapter showed the properties required for the Tweeter WFS in ExAO: high efficiency, high-speed capabilities, high spatial resolution, and a large dynamic range. Focusing on these requirements, this chapter reviews and classifies the representative WFSs for AO that have been demonstrated to date.

4.1 Types of wavefront sensors and their efficiency

This section introduces a classification of WFSs, which convert the wavefront aberration into light intensity distributions (signals). Depending on the relationship between the wavefront aberration and the signals, WFSs can be classified into three types: slope, curvature, or phase sensors. Depending on this classification, the frequency characteristics of the efficiency is different. The reconstruction methods to calculate the wavefront aberration from the signals can be classified into zonal or modal, which set different efficiency measures.

4.1.1 Relationship between wavefront aberrations and signals

The coordinate and the phase aberration are defined for the classification. The position of a subaperture is (x, y) in the telescope aperture. The wavefront aberration in the subaperture is $W(x, y)$ [nm]. The phase aberration $\delta(x, y)$ [rad] is defined as:

$$\delta(x, y) = \frac{2\pi}{\lambda_C} W(x, y), \tag{4.1}$$

where λ_C [nm] is the center of the wavelength band used by the WFS. Depending on the relationship between $\delta(x, y)$ and the signals, WFS is classified into three types.

- **Curvature sensors**: The signals reflect the local curvature $\nabla^2\delta(x, y)$.
- **Slope sensors**: The signals reflect the local slope $\nabla\delta(x, y)$.
- **Phase sensors**: The signals reflect the phase aberration $\delta(x, y)$ itself when the aberration is small.

These classes have different frequency characteristics of efficiency. The curvature and slope sensors have low efficiency for phase aberrations with low spatial frequencies. Such aberrations have loose slopes or curvatures, which produce weak signals and result in low S/N ratios and significant measurement errors. On the other hand, the phase sensors, in principle, have a constant efficiency regardless of spatial frequency.

These frequency characteristics have impacts on the contrast after ExAO. The spatial frequency of the wavefront aberration determines the position of the speckles, as shown in Section 2.1.3. A low frequency corresponds to a small angular separation from the host star. Thus, the curvature/slope sensors have low sensitivity to speckles near the host star. On the other hand, the phase sensors have constant and relatively high sensitivity to speckles regardless of their angular separations [27]. Therefore, the phase sensors are preferable for the Tweeter WFS in ExAO that suppresses speckles near the host star.

4.1.2 Reconstruction

Reconstruction stands for the calculation of the phase aberration from the signals. There are two types of reconstructions.

- **Zonal reconstruction**: The phase, slope, or curvature at each subaperture (x, y) is determined only from the signals corresponding to the subaperture. The curvature/slope sensors require a further 2-D integration of the measured curvature/slope to calculate the phase aberration.
- **Modal reconstruction**: The phase aberration is determined without calculating the phase, slope, or curvature at each subaperture. The modal reconstruction expresses the phase aberration $\delta(x, y)$ as a linear

combination of the known bases $\phi_m(x, y)$:

$$\delta(x, y) = \sum_{m=1}^{M} A_m \phi_m(x, y), \tag{4.2}$$

where A_m is the amplitude of each basis, and M ($\leq N_{SA}$) is the total number of the bases. Examples of the bases are the Fourier modes and the Zernike modes. The amplitude A_m is calculated from the WFS signals. In the calculation, the captured signal $S_\delta(x, y)$ of the unknown phase $\delta(x, y)$ is approximated as a linear combination of the known signals $S_{\phi m}(x, y)$ of $\phi_m(x, y)$:

$$S_\delta(x, y) = \sum_{m=1}^{M} A_m S_m(x, y). \tag{4.3}$$

Based on the captured signal $S_\delta(x, y)$ and the known signals $S_m(x, y)$, the unknown amplitudes A_m is calculated with several methods such as a matrix operation.

These reconstructions have different efficiency measures. Note that only photon noise is considered here as the noise that produces measurement errors of the WFS; readout noise is not considered.

- C_p **for zonal reconstruction**: The RMS error σ_{pn} [rad] of the phase-aberration measurement is expressed as:

$$\sigma_{pn} = \frac{C_p}{\sqrt{N_e}}, \tag{4.4}$$

where C_p is the coefficient of error propagation, and N_e is the average number of photons (in the number of photo-electrons) incident in a subaperture. A small value of C_p represents a high efficiency.

- β_p **for modal reconstruction**: The sensitivity factor β_p [27] is defined when the bases are the Fourier modes. A sinusoidal phase aberration with a singe spatial frequency is assumed. The WFS measures the amplitudes A_{sin}/A_{cos} of the sine/cosine components and reconstructs the phase aberration. The RMS error Σ [rad] of the reconstructed phase aberration is expressed as:

$$\Sigma = \frac{\beta_p}{\sqrt{N_e \times N_{SA}}}, \tag{4.5}$$

where N_{SA} is the number of the subapertures used in the measurement. A small value of β_p represents a high efficiency. The minimum $\beta_p = 1$.

4.2 Examples of curvature sensors

There are various examples of Curvature WFS (CWFS) optical systems [37, 38, 39]. This section introduces the CWFS by Roddier [37].

The optical system consists of a convex lens L_1 with a focal length of f and another convex lens L_2 with a focal length of $f/2$ placed at the focal plane of L_1 (Fig. 4.1(a)). The signals $I_1(x, y)$ and $I_2(x, y)$ are respectively captured at the two surfaces P_1 and P_2, separated from L_2 by a distance l back and forth. L_1 makes P_1 conjugated to the surface P_1', which is separated from the wavefront W by a distance $(f - l)f/l$ forward. Also, L_1 and L_2 makes P_2 conjugated to the surface P_2', which is separated from the wavefront W by a distance $(f - l)f/l$ backward (Fig. 4.1(b)). If the wavefront W is convex with respect to the direction of propagation, the rays diverge from P_2' to P_1', creating bright and dark regions on P_2', P_1', respectively. Thus, the difference between I_2 and I_1 reflects the local curvature of the wavefront W. The relationship is formulated as follows:

$$\frac{I_2(x,y) - I_1(x,y)}{I_2(x,y) + I_1(x,y)} = \frac{f(f-l)}{2l} \left\{ \nabla^2 W(x,y) + \frac{\partial}{\partial n} W(x,y) \delta_C \right\}, \tag{4.6}$$

where $\partial/\partial n$ represents the derivative in the radial direction of the beam. δ_C is defined as follows:

$$\delta_C = \begin{cases} 0 & \text{(when } (x,y) \text{ is inside of the beam,)} \\ \infty & \text{(when } (x,y) \text{ is at the edge of the beam.)} \end{cases} \tag{4.7}$$

These equations enable the zonal reconstruction.

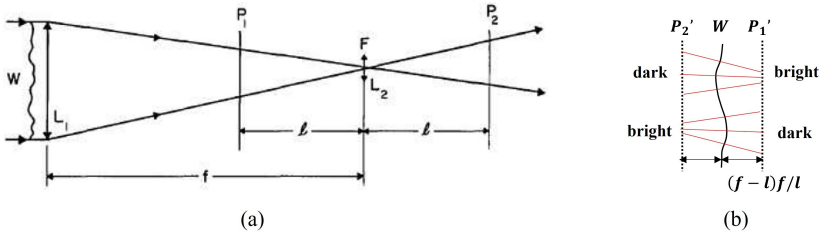


Figure 4.1: (a) The CWFS optical system, cited from Roddier. [37] The black double arrows L_1 and L_2 represent the convex lenses. The black single arrows represent the edge of the beam. (b) Location of the conjugate surfaces P_1' and P_2' of P_1 and P_2 relative to the wavefront W. Solid red lines represent rays.

4.3 Examples of slope sensors

4.3.1 Shack-Hartmann wavefront sensor (SHWFS)

The SHWFS [40] consists of a micro-lens array (MLA) and a detector at the focal plane of the MLA (Fig. 4.2). The MLA creates a focal spot array on the detector. One element of the MLA corresponds to one subaperture. The position of the spot generated by each element is related to the local wavefront slope $\nabla W(x, y)$ at each subaperture (x, y). In the coordinate system on the detector, the coordinate (X, Y) [mm] represents the position of the center of gravity (CoG) of the spot corresponding to one subaperture. The coordinate (X_0, Y_0) [mm] represents the position of the CoG of the spot relative to the reference plane wave. The slope of the wavefront at that subaperture is given by

$$\nabla W(x, y) = \frac{1}{F}(X - X_0,\ Y - Y_0), \tag{4.8}$$

where F [mm] is the focal length of the MLA. This equation enables the zonal reconstruction.

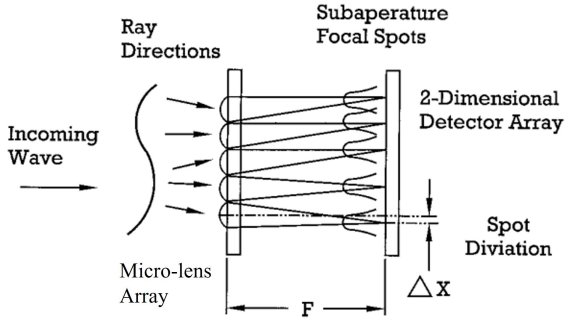


Figure 4.2: The conseptual drawing of the SHWFS, cited from Platt & Shack. [40] The spot deviation $\Delta X = X - X_0$.

4.3.2 Modulated pyramid wavefront sensor (MPyWFS)

The MPyWFS, originally proposed by Ragazzoni [41], consists of a tip/tilt mirror, a pyramid prism, a converging lens, a relay lens, and a detector (Fig. 4.3(a)). The tip/tilt mirror is a flat mirror whose normal direction is variable. The pyramid prism is located at the focal plane of the converging

lens. The detector is placed at the output pupil. In the following, the principle with geometrical optics is explained using a ray from the subaperture A (position (x, y)). The tip/tilt mirror periodically modulates the direction of the ray. The intersection of the ray and the focal plane moves in a circle with a radius of α [rad] at a constant speed. The circle is called the modulation path. The ray is refracted with the pyramid prism to intersect the detector at any one point of the four points $(A_1 - A_4)$; each of the four points corresponds to one of the four quadrants of the pyramid. The detector is exposed during the modulation period and captures the signals $I_1(x, y) - I_4(x, y)$. When the local wavefront slope is zero at the subaperture A, the modulation path's center matches the pyramid's apex. Thus, the ray from A stays in the four quadrants for the same amount of time each, producing the equal signals $I_1(x, y) - I_4(x, y)$. When the local slope $\partial W/\partial x > 0$ at A, the modulation path's center moves by ΔX [rad] in the x direction (Fig. 4.3(b)):

$$\Delta X = \frac{\partial W(x, y)}{\partial x}. \tag{4.9}$$

Thus, the ray from A stays in the two quadrants at $x > 0$ longer than the other two quadrants, producing larger signals $I_1(x, y)$ and $I_4(x, y)$ than $I_2(x, y)$ and $I_3(x, y)$. Therefore, the difference between $[I_1(x, y) + I_4(x, y)]$ and $[I_2(x, y) + I_3(x, y)]$ reflects the local wavefront slope in the x direction. The following equation describes the relationship between the signals and the local slope of the phase aberration $\delta(x, y)$ [42, 43]:

$$\frac{\partial \delta(x, y)}{\partial x} = \frac{\alpha \pi^2}{\lambda_C} \sin(S_x(x, y)) \tag{4.10}$$

$$S_x(x, y) = 4 \frac{[I_1(x, y) + I_4(x, y)] - [I_2(x, y) + I_3(x, y)]}{[I_1(x, y) + I_4(x, y)] + [I_2(x, y) + I_3(x, y)]}, \tag{4.11}$$

where $S_x(x, y)$ is the normalized signal. The equation enables the zonal reconstruction.

According to analysis with wave optics [42], the MPyWFS acts as a slope sensor with the low-order phase aberrations whose spatial frequencies are $< \alpha/\lambda_C$. The MPyWFS also acts as a phase sensor with higher-order aberrations; the principle is similar to the fixed pyramid WFS described in the next section.

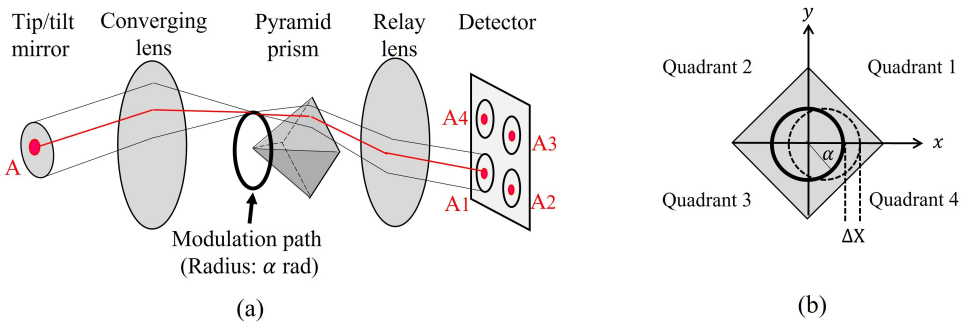


Figure 4.3: The conceptual drawings of the MPyWFS. (a) Overview. The thickness of the pyramid prism in the optical axis direction is depicted with emphasis. (b) The pyramid prism viewed from the incident side, showing the four quadrants and the modulation paths.

4.4 Examples of phase sensors

4.4.1 Fixed pyramid wavefront sensor (FPyWFS)

The FPyWFS, also called non-modulated PyWFS, consists of a pyramid prism, a converging lens L_1, a relay lens L_2, and a detector (Fig 4.4(a)). The pyramid prism is located at the focal plane of the converging lens. The detector is placed at the output pupil. The FPyWFS does not include the tip/tilt mirror; the PSF core is fixed at the pyramid prism's apex. The pyramid prism divides the PSF into four parts, which produce the four signals $I_1(x,y) - I_4(x,y)$ on the detector. These signals are interpreted as the interferograms of the divided PSF core and the speckles at each quadrant [17].

Using a diffraction theory, the relationship between the phase aberration $\delta(x,y)$ and the signals are described as follows [44, 45]:

$$S_x(x,y) = \frac{[I_1(x,y) + I_4(x,y)] - [I_2(x,y) + I_3(x,y)]}{|u_0|^2 (f_1/f_2)^2}, \tag{4.12}$$

$$\begin{aligned} S_x(x,y) = &\int_{-P(y)}^{P(y)} dx' \, \frac{\sin[\delta(x,y) - \delta(x',y')]}{2\pi(x-x')} \\ &+ \int_{-P(x)}^{P(x)} dy_2 \int_{-P(x)}^{P(x)} dy_1 \int_{-P(y_1)}^{P(y_1)} dx_1 \frac{\sin[\delta(x,y_2) - \delta(x_1,y_1)]}{2\pi^3 (x-x_1)(y-y_1)(y-y_2)}. \end{aligned} \tag{4.13}$$

In the equations, $S_x(x,y)$ is the normalized signal. $|u_0|^2$ is the average

intensity of the incident light at each subaperture. f_1 and f_2 are the focal lengths of L_1 and L_2, respectively. $P(x)/P(y)$ denote the y/x coordinates of the points at the edge of the entrance pupil; the points have the same x/y coordinates as the subaperture (x, y) (Fig. 4.4(b)).

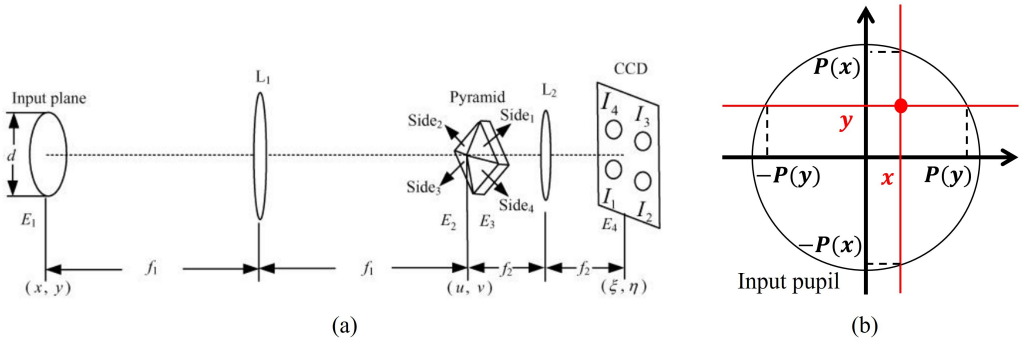


Figure 4.4: (a) The conceptual drawing of the FPyWFS, cited from Wang et al. [44] (b) The definition of the coordinates $P(x)$ and $P(y)$ in Eq. 4.13

The FPyWFS has various methods for the reconstruction [43]. Among them, a standard modal reconstruction called MVM (matrix vector multiplication) is described below. This method is widely used in actual facilities. When the phase aberration $|\delta| \ll 1$ rad, Eq. 4.13 is linearized with the following approximations:

$$\begin{aligned} \sin[\delta(x, y) - \delta(x', y')] &= \delta(x, y) - \delta(x', y'), \\ \sin[\delta(x, y_2) - \delta(x_1, y_1)] &= \delta(x, y_2) - \delta(x_1, y_1). \end{aligned} \tag{4.14}$$

In addition, δ is expanded with the known bases ϕ_m:

$$\delta(x, y) = \sum_{m=1}^{N_{SA}} A_m \phi_m(x, y), \tag{4.15}$$

where A_m is the unknown amplitude. The approximations and expansion result in the following linear approximation of Eq. 4.13:

$$S_x(x, y) = \sum_{m=1}^{N_{SA}} A_m S_m^x(x, y). \tag{4.16}$$

In this approximation, S_m^x is the known normalized signal of the m-th mode

ϕ_m. Similarly, the normalized signal $S_y(x, y)$ is linearized as:

$$S_y(x, y) = \frac{[I_1(x, y) + I_2(x, y)] - [I_3(x, y) + I_4(x, y)]}{|u_0|^2 (f_1/f_2)^2}, \tag{4.17}$$

$$S_y(x, y) = \sum_{m=1}^{N_{SA}} A_m S_m^y(x, y), \tag{4.18}$$

where S_m^y is the known normalized signal of the m-th mode ϕ_m. $S_x(x, y)$ and $S_y(x, y)$ are measured at all the N_{SA} subapertures; the number of measured value is $2N_{SA}$. These values are aligned in one dimension and defined as the vector $\mathbf{S}$. Similarly, the N_{SA} values of amplitude A_m are aligned and defined as the vector $\mathbf{A}$. Using the Eqs. 4.16 and 4.18, the vectors $\mathbf{S}$ and $\mathbf{A}$ is related with the known matrix T as:

$$\mathbf{S} = T\mathbf{A}, \tag{4.19}$$

The components of the matrix T are $S_m^x(x, y)$ and $S_m^y(x, y)$. Thus, the unknown $\mathbf{A}$ is calculated through multiplying the measured $\mathbf{S}$ by the known pseudo-inverse matrix T^{-1}:

$$\mathbf{A} = T^{-1}\mathbf{S} \tag{4.20}$$

The calculated $\mathbf{A}$ determines the phase aberration $\delta(x, y)$ using Eq. 4.15.

4.4.2 Zernike wavefront sensor (ZWFS)

The ZWFS, originally proposed by Zernike [46], is a kind of the PDI (Point-Diffraction interferometer). The PDI (Fig. 4.5) includes a focusing lens and a point-diffraction plate placed at the focal plane. The plate is transparent and has a pinhole smaller than the PSF core. The pinhole extracts an unaberrated beam (reference beam) from the incident beam under test. The rest of the beam (test beam) retains the phase information of the incident beam. The two beams interfere, and an interferogram arises on a detector. The intensity distribution of the interferogram reflects the phase aberration of the incident beam.

The ZWFS has the following properties. The pinhole radius is $\simeq F\lambda_C$, where F is the F-number of the focused beam onto the point-diffraction plate. The phase shift of the reference beam induced by the pinhole is $+\pi/2$ rad or $\pm\pi/2$ rad; each type is described as follows.

A. $+\pi/2$ type

The practical example is ZELDA (Zernike sensor for Extremely Low-level Differential Aberrations) [48, 49]. The signal I_C against the phase

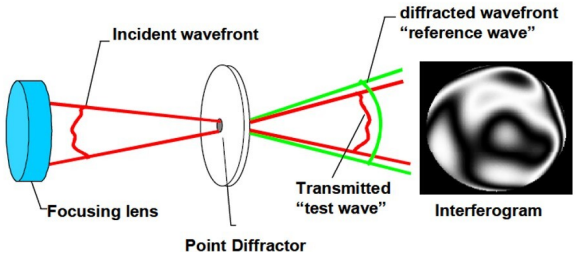


Figure 4.5: The conceptual drawing of the PDI, cited from Millerd et al. [47] The reference and test wave indicate the wavefront of the reference and test beam, respectively.

aberration δ is formulated as:

$$I_C = P^2 + 2b^2 + 2Pb(\sin\delta - \cos\delta) \tag{4.21}$$

$$= P^2 + 2b^2 + 2\sqrt{2}Pb\sin(\delta - \pi/4), \tag{4.22}$$

where P is the amplitude of the incident beam, and b ($\simeq 0.5P$) is the amplitude of the reference beam. Normalizing to $P = 1$ and assuming b is known, δ is calculated as follows:

$$\delta = \arcsin\left(\frac{I_C - 1 - 2b^2}{2\sqrt{2}b}\right) + \frac{\pi}{4}. \tag{4.23}$$

An approximate solution of δ is calculated by expanding Eq. 4.21 [48]:

$$I_C = 1 + 2b^2 + 2b(\delta^2/2 + \delta - 1), \tag{4.24}$$

in which P is normalized to unity. Solving Eq. 4.24 yields the approximated δ as follows:

$$\delta = -1 + \sqrt{3 - 2b - (1 - I_C)/b}. \tag{4.25}$$

These equations enable the zonal reconstruction.

B. $\pm\pi/2$ type

The practical example is the vector-Zernike WFS [50]. The signals I_R and I_L with phase shifts of $\pm\pi/2$ are formulated as:

$$I_R = \frac{1}{2}\left\{P^2 + 2b^2 + 2Pb(\sin\delta - \cos\delta)\right\}, \tag{4.26}$$

$$I_L = \frac{1}{2}\left\{P^2 + 2b^2 - 2Pb(\sin\delta + \cos\delta)\right\}. \tag{4.27}$$

δ is calculated with the differential of these signals:

$$\delta = \arcsin\left(\frac{I_R - I_L}{2Pb}\right), \tag{4.28}$$

$$P = \sqrt{I_R + I_L + \sqrt{4b^2(I_R + I_L) - (I_R - I_L)^2 - 4b^2}}. \tag{4.29}$$

These equations enable the zonal reconstruction.

4.5 Comparison of existing wavefront sensors

This section compares the WFSs exemplified above. This section focuses on the requirements presented in Section 3.4 except for "high spatial resolution" because the WFSs can achieve high spatial resolution by increasing the number of detector pixels or micro lenses.

4.5.1 Efficiency

Guyon calculated the sensitivity factor β_p of each WFS to compare the efficiency [27]. Note that β_p depends on the spatial frequency of the wavefront aberration, i.e., the speckle separation angle. The Curvature/slope sensors have relatively large β_p ($\gtrsim 10$) for low spatial frequency wavefronts. In other words, the efficiency is low for speckles in the vicinity of the host star. On the other hand, the phase sensor has a constant β_p ($\lesssim 2$) regardless of the spatial frequency; the efficiency does not decrease even for speckles near the host star. Therefore, the phase sensors are suitable for ExAO. Figure 4.6 summarizes the above characteristics.

4.5.2 High-speed capabilities

- **Calculation cost**: The total number of the subapertures is N_{SA}. The calculation cost of each WFS is summarized in Table 4.1. The curvature/slope sensors require the calculation of the curvature/slope at each subaperture and 2-D integration of them. The former has a calculation cost of $\mathcal{O}(N_{SA})$; the latter has a cost of $\mathcal{O}(N_{SA}^2)$ because it requires a matrix operation with $\sim N_{SA} \times N_{SA}$ elements. The latter cost surpasses the former in the limit where N_{SA} is large. The phase sensors have various calculation costs. The FPyWFS have various reconstruction methods, whose costs range $\mathcal{O}(N_{SA}) - \mathcal{O}(N_{SA}^2)$ [43]. The standard method MVM described in Section 4.4.1 has a cost of $\mathcal{O}(N_{SA}^2)$ because it uses a matrix with $\sim N_{SA} \times N_{SA}$ elements.

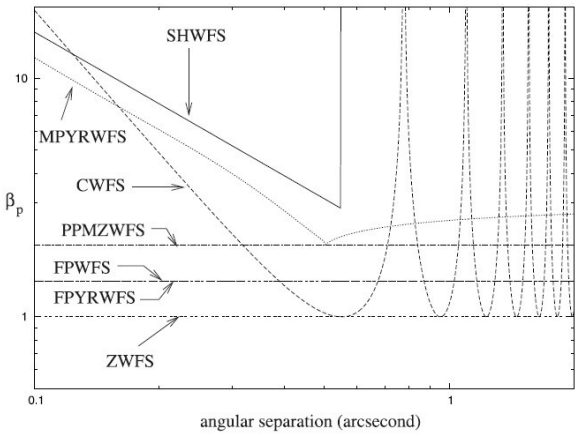


Figure 4.6: The comparison of β_p, cited from Guyon. [27] The horizontal axis is the angular separation from the host star, which is converted from the spatial frequency under $D = 8$ m and the imaging wavelength $\lambda_i = 1.6$ μm. MPYRWFS: MPyWFS. FPYRWFS: FPyWFS. PPMZWFS: Pupil-plane Mach-Zehnder WFS. FPWFS: Focal plane WFS.

It has been noted that such a calculation is challenging for ExAO with $N_{SA} \sim 40000$ [43]. On the other hand, the ZWFS has a cost of $\mathcal{O}(N_{SA})$ because the phase aberration δ is directly calculated at each subaperture.

- **Size of the readout region**: The 1-D number of the subapertures is N_{1d}. The size of the readout region of each WFS is summarized in Table 4.1. The CWFS and the ZWFS produces one or two signals on the detector, whose one pixel corresponds to one subaperture. Thus, the vertical size of the readout region is equal to N_{1d} pixels. The SHWFS requires at least 2×2 pixels to detect the CoG of each spot. Thus, the vertical size of the readout region is $\geq 2N_{1d}$ pixels to deal with vertical N_{1d} spots (subapertures). The MPyWFS and FPyWFS produces 2×2 signals on the detector, whose one pixel corresponds to one subaperture. Each signal also has a finite spacing between the adjacent signals. Thus, the vertical size of the readout region is $\geq 2N_{1d}$ pixels.

4.5.3 Dynamic range

This section compares each WFS's dynamic range characterized by the range in which the reconstruction equation is applicable. The curvature/slope sensors' dynamic ranges can extend to $> 1\lambda_C$. On the other hand, the phase sensors have relatively small dynamic ranges. In the case of the FPyWFS, the approximation of sine (Eq. 4.14) is applicable only when $|\delta| < \pi/2$, i.e., peak-to-valley (P-V) value is less than $\lambda_C/2$ [44]. Thus, the dynamic range is $\lesssim \lambda_C/2$. Larger aberrations cause ambiguity. In the case of the ZWFS, the proposed equations calculate δ from the signal I_C or the differential signal $I_R - I_L$. The signal and the differential signal reflect the sine of the phase aberration δ. Thus, the equations are applicable in $|\delta| < \pi/2$ without causing ambiguity; the dynamic range is $< \lambda_C/2$.

4.6 Summary of this chapter

The comparison in this chapter is summarized in Table 4.1. Regarding efficiency, ExAO favors the phase sensors such as the FPyWFS and ZWFS. However, the FPyWFS has a relatively high calculation cost and readout region, which are disadvantageous for high-speed measurement. The ZWFS has a lower calculation cost and a smaller readout region but has a small dynamic range. Therefore, there is room for improvement in the dynamic range and high-speed capabilities of phase sensors.

Table 4.1: Comparison of existing WFSs. C_p: coefficient of error propagation from photon noise. β_p: sensitivity factor [27]. λ_C: center of the sensing wavelength band. N_{SA}: number of all the effective subapertures. N_{1d}: 1-dimensional number of the effective subapertures.

	CWFS	SHWFS	MPyWFS
Type	Curvature	Slope	Slope
C_p	N/A	N/A	N/A
β_p	≥ 1	> 2	≥ 2
Calculation cost	$\mathcal{O}(N_{SA}^2)$	$\mathcal{O}(N_{SA}^2)$	$\mathcal{O}(N_{SA}^2)$
Number of detector pixels (vertical)	N_{1d}	$\geq 2N_{1d}$	$> 2N_{1d}$
Dynamic range	$> \lambda_C$ [ii]	$> \lambda_C$ [ii]	$> \lambda_C$ [ii]

	ZWFS	FPyWFS
Type	Phase	Phase
C_p	$\sqrt{0.5}$ [i]	N/A
β_p	1.0	$\sqrt{2}$
Calculation cost	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA}) - \mathcal{O}(N_{SA}^2)$
Number of detector pixels (vertical)	N_{1d}	$> 2N_{1d}$
Dynamic range	$< 0.5\lambda_C$ [ii]	$\lesssim 0.5\lambda_C$ [ii]

[i] From N'Diaye et al. 2013 [48].

[ii] The range in which the reconstruction equation is applicable.

Chapter 5

Principle of the New Wavefront Sensor

This chapter describes a new phase sensor with a large dynamic range and high-speed capabilities. The new sensor is named b-PDI (birefringent point-diffraction interferometer). Table 5.1 shows the target performance. The b-PDI is based on the PDI (Fig.4.5) to obtain a high efficiency and high-speed capabilities. In addition, the b-PDI utilizes the phase-shifting interferometry to obtain a large dynamic range. This method is applicable in P-V $< 1\lambda_C$ without causing ambiguity because it calculates the tangent (sine/cosine) of the phase aberration.

This chapter is based on Sections 2 and 3 of Tsukui et al. [51], published in Applied Optics.

Table 5.1: Target performance of the new phase sensor. β_p: sensitivity factor [27]. N_{1d}: 1-dimensional number of the effective subapertures. λ_C: center of the sensing wavelength band.

Type	Phase sensor
β_p	≤ 1.4
Sampling rate	6.5 kHz [i]
Calculation cost	$\mathcal{O}(N_{SA})$
Number of detector pixels (vertical)	N_{1d}
N_{1d}	≥ 24 [i]
Dynamic range	$1\lambda_C$
λ_C	800 nm [i]

[i] Assuming SEICA instrument [17].

5.1 Principles

This section describes the principles of the b-PDI and formulate a reconstruction algorithm.

5.1.1 Overview

The b-PDI utilizes a point-diffraction plate made of birefringent crystal (Fig. 5.1). The substrate of the plate (hereafter birefringent point-diffraction beam-splitter; BPBS) is made of uniaxial birefringent crystal. Its optical axis is placed parallel to the y-axis in Fig 5.1. With this configuration, the refractive indices of the crystal are n_o and n_e for linearly polarized beam parallel to the x and y-axes, respectively. The BPBS has a pinhole of the radius R_p and depth d on the front surface. The pinhole is filled with non-birefringent material, whose refractive index is n. The ideal n satisfies the equation

$$n = \frac{n_e + n_o}{2}. \tag{5.1}$$

The thickness of the layer of non-birefringent material is the same as the depth of the pinhole. Thus, the front surface of the BPBS is flat. The width of the clear aperture of the BPBS is l, which is defined later.

The incident beam is focused on the front surface of the BPBS to form an Airy disk. The diameter of the Airy disk is set to be larger than the pinhole.

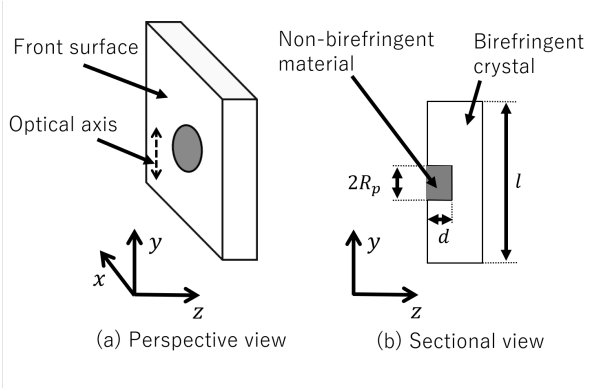


Figure 5.1: Schematic drawing of the BPBS. (a) Perspective view. The optical axis of the birefringent crystal is parallel to the y-axis. (b) Sectional view. the BPBS has a pinhole with the radius R_p and depth d on the front surface. The pinhole is filled with non-birefringent material. The width of the clear aperture is l. Reprinted with permission from [51] ©Optica Publishing Group.

In the transmission path, the beam passing through the inner region of the pinhole is interpreted as a reference beam with an ideal spherical wavefront, while the beam passing through the outer region keeps the wavefront information of the incident beam and is interpreted as a test beam. The test and the reference beams interfere with each other and produce an interferogram. The optical path difference (OPD) between the two beams from the inner and outer regions gives the phase difference between the reference and test beams. The OPD differs depending on the polarization direction due to birefringence. The phase differences θ_o and θ_e are defined according to polarization components along the x and y-axes, respectively:

$$\theta_o = \frac{2\pi}{\lambda_C}(n - n_o)d = \frac{\pi d(n_e - n_o)}{\lambda_C} \tag{5.2}$$

$$\theta_e = \frac{2\pi}{\lambda_C}(n - n_e)d = -\frac{\pi d(n_e - n_o)}{\lambda_C}, \tag{5.3}$$

where λ_C denotes the wavelength of the incident beam. We set $\lambda_C = 800$ nm in our design. Choosing d as of

$$d = \frac{2\lambda_C}{3(n_e - n_o)} \tag{5.4}$$

yields $\theta_o = 2\pi/3$ rad and $\theta_e = -2\pi/3$ rad. In the reflection path, the reference and test beams are generated with the same manner as the transmitted

beam. The phase difference between them is 0 rad because the front surface of the BPBS is flat. To summarize, the BPBS simultaneously provides interferograms of phase differences of 0 and $\pm 2\pi/3$ rad. The system utilizes all the incident photons except for those diffracted by the pinhole at large angles and falling outside the interferograms.

The incident beam is split into four interferograms with the optical system shown in Fig. 5.2. The optical system consists of a focusing lens L0, BPBS, collimator lenses L1 and L2, Savart plates SP1 and SP2, and detectors D1 and D2. The BPBS is placed at the focal plane of L0. The axis of L0 aligns with the center of the pinhole on the BPBS. The angle between the axis of L0 and the normal of the BPBS is 10 deg. L1, L2, and L0 share the common focal point. L1 and SP1 form a pair of the exit pupils on D1. L2 and SP2 form another pair of the exit pupils on D2. The incident beam is focused with L0 to form an Airy disk of F/31 on the front surface of the BPBS. The radius of the Airy disk R_{Airy} is 30 μm when the wavelength λ_C is 800 nm. The transmitted beam through the BPBS is then collimated with L1 and split into two orthogonal polarization components with SP1; one component is parallel to the x-axis and forms interferogram $I_o^T(x, y)$ with a phase difference of $\theta_o = 2\pi/3$ rad on D1 whereas the other is parallel to the y-axis and forms interferogram $I_e^T(x, y)$ with a phase difference of $\theta_e = -2\pi/3$ rad on D1. The reflected beam on the BPBS is similar; collimated with L2 and split into two with SP2: one parallel to the x-axis forming interferogram $I_o^R(x, y)$ on D2 and the other parallel to the y-axis forming interferogram $I_e^R(x, y)$ on D2. Both of them are interferograms with a phase difference of 0 rad. This configuration makes it easier to formulate the reconstruction algorithm. These four interferograms are simultaneously captured with D1 and D2. The diameter of each interferogram is N_{1d} pixels on D1 and D2. Two interferograms line up horizontally on each detector. Thus, the size of the readout dregion is N_{1d} in the vertical direction.

We set the half width $l/2$ of the clear aperture as a function of N_{1d}:

$$\frac{l}{2} = F\lambda_C \times \frac{N_{1d}}{2}, \tag{5.5}$$

where F is the F-number of the focused beam with L0. The BPBS with $l/2$ acts as a spatial filter to filter out wavefront aberrations with higher spatial frequencies than the Nyquist frequency. This spatial filtering smooths the interferograms and reduces the effect of position mismatch between them. When $N_{1d} = 24$ and $F = 31$, l is set to 0.6 mm.

In the design, the incident angle of the beam to BPBS is 10 deg. Concerns might be raised about potentially serious effects of such an oblique incidence

to the birefringent material on the wavefront shape. However, Mercer et al. [52] successfully measured the wavefront with a point-diffraction plate made of birefringent liquid-crystal tilted by as large as 20 deg. Their result suggests that the oblique incidence in our design should be acceptable.

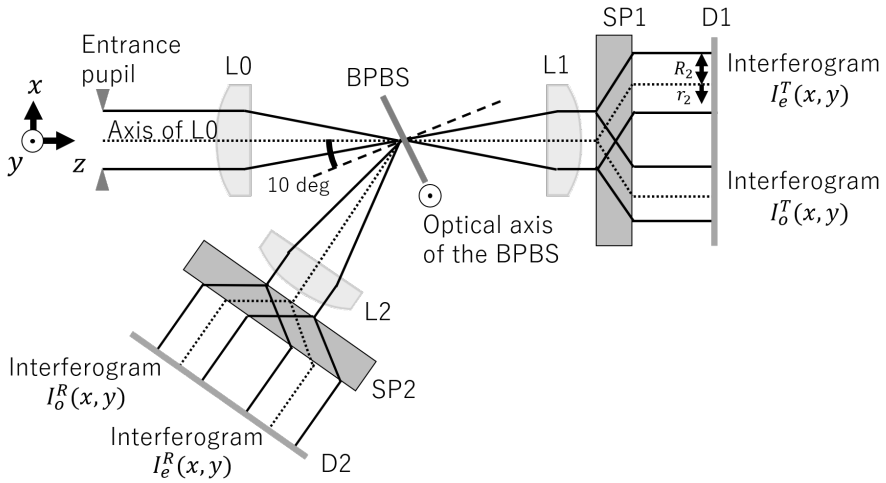


Figure 5.2: Schematic drawing of the optical system of b-PDI. The angle between the normal of the BPBS (dashed line) and the axis of L0 is 10 deg. Solid and dotted lines represent outer edges and the axes of the beams, respectively. Reprinted with permission from [51] ©Optica Publishing Group.

5.1.2 Formulation of the four interferograms

We analytically formulate a reconstruction algorithm. In this section, I formulate the intensity of the four interferograms. Firstly, the complex amplitude of the incident beam on the input pupil is given by the following Jones vector:

$$\boldsymbol{E}_1(x,y) = \begin{pmatrix} E_o \\ E_e \end{pmatrix} \{1+\alpha(x,y)\} \exp\{i\delta(x,y)\}, \tag{5.6}$$

where E_o and E_e are the mean amplitudes of the incident beam along the x and y-axes, respectively, (x,y) is the coordinates of the input pupil, $\alpha(x,y)$ is the deviation of the amplitude at point (x,y) from the mean amplitude, and phase aberration $\delta(x,y)$ is the deviation of the phase at point (x,y) from the uniform phase distribution. Secondly, the magnification between the input and exit pupils is 1; this means that L0, L1, and L2 have a common focal length. Finally, the polarization properties of the BPBS are given by the following Jones matrices, following the expressions in Imada et al. [53] (hereafter I15).

$$T(r_B) = \begin{cases} \begin{pmatrix} T^o_{\text{in}} \exp(2i\pi/3) & 0 \\ 0 & T^e_{\text{in}} \exp(-2i\pi/3) \end{pmatrix} & 0 \le r_B \le R_p \\ \begin{pmatrix} T^o_{\text{out}} & 0 \\ 0 & T^e_{\text{out}} \end{pmatrix} & R_p < r_B \le R_B \end{cases} \tag{5.7}$$

$$R(r_B) = \begin{cases} \begin{pmatrix} R^o_{\text{in}} & 0 \\ 0 & R^e_{\text{in}} \end{pmatrix} & 0 \le r_B \le R_p \\ \begin{pmatrix} R^o_{\text{out}} & 0 \\ 0 & R^e_{\text{out}} \end{pmatrix} & R_p < r_B \le R_B, \end{cases} \tag{5.8}$$

where $T(r_B)$ and $R(r_B)$ are the transmission and reflection properties, respectively, r_B is the distance from the center of the pinhole, and R_p and R_B are the radii of the pinhole and the BPBS, respectively. T^b_a and R^b_a denote the amplitude-transmission coefficient and the amplitude-reflection coefficient, respectively, where the subscript a indicates the inside (in) or the outside (out) of the pinhole, and the superscript b indicates the polarizing direction parallel to the x-axis (o) or parallel to the y-axis (e).

Using Eqs. 21 and 22 in I15, the complex amplitudes of interferograms

$I_o^T,\ I_e^T,\ I_o^R$, and I_e^R are respectively described as,

$$E_{2,o}^T(x,y) = -E_o\left\{[T_{\text{in}}^o \exp(2i\pi/3) - T_{\text{out}}^o]f + T_{\text{out}}^o(1+\alpha)\exp(i\delta)\right\} \tag{5.9}$$

$$E_{2,e}^T(x,y) = -E_e\left\{[T_{\text{in}}^e \exp(-2i\pi/3) - T_{\text{out}}^e]f + T_{\text{out}}^e(1+\alpha)\exp(i\delta)\right\} \tag{5.10}$$

$$E_{2,o}^R(x,y) = -E_o\left\{(R_{\text{in}}^o - R_{\text{out}}^o)f + R_{\text{out}}^o(1+\alpha)\exp(i\delta)\right\} \tag{5.11}$$

$$E_{2,e}^R(x,y) = -E_e\left\{(R_{\text{in}}^e - R_{\text{out}}^e)f + R_{\text{out}}^e(1+\alpha)\exp(i\delta)\right\}, \tag{5.12}$$

where f is the normalized amplitude of the reference beam and is a function of R_p and r_2, the latter of which is the distance from the center of each exit pupil (Fig. 5.2). The function $f(R_p; r_2)$ is specified analytically by I15 and shown in Fig. 5.3.

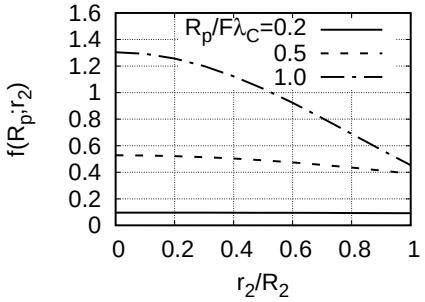


Figure 5.3: Function $f(R_p; r_2)$ based on Eq. 17 in I15. The horizontal axis is the distance r_2 from the center of an exit pupil normalized by the radius of the pupil R_2 (see Fig. 5.2). The three curves correspond to $R_p/F\lambda_C = 0.2, 0.5, 1.0$; F is the F-number of the focused beam with L0. Reprinted with permission from [51] ©Optica Publishing Group.

The new parameters $\tau,\ \phi$, and ρ are defined as follows.

$$\tau_o \exp(i\phi_o) = \frac{T_{\text{in}}^o}{T_{\text{out}}^o}\exp(2i\pi/3) - 1 \tag{5.13}$$

$$\tau_e \exp(i\phi_e) = \frac{T_{\text{in}}^e}{T_{\text{out}}^e}\exp(-2i\pi/3) - 1 \tag{5.14}$$

$$\rho_o = \frac{R_{\text{in}}^o}{R_{\text{out}}^o} - 1 \tag{5.15}$$

$$\rho_e = \frac{R_{\text{in}}^e}{R_{\text{out}}^e} - 1 \tag{5.16}$$

The intensity maps of each interferogram are given as follows.

$$I_o^T(x,y) = (E_o T_{\text{out}}^o)^2 \left\{\tau_o^2 f^2 + (1+\alpha)^2 + 2\tau_o f(1+\alpha)\cos(\phi_o - \delta)\right\} \quad (5.17)$$

$$I_e^T(x,y) = (E_e T_{\text{out}}^e)^2 \left\{\tau_e^2 f^2 + (1+\alpha)^2 + 2\tau_e f(1+\alpha)\cos(\phi_e - \delta)\right\} \quad (5.18)$$

$$I_o^R(x,y) = (E_o R_{\text{out}}^o)^2 \left\{\rho_o^2 f^2 + (1+\alpha)^2 + 2\rho_o f(1+\alpha)\cos\delta\right\} \quad (5.19)$$

$$I_e^R(x,y) = (E_e R_{\text{out}}^e)^2 \left\{\rho_e^2 f^2 + (1+\alpha)^2 + 2\rho_e f(1+\alpha)\cos\delta\right\} \quad (5.20)$$

The absolute values of the amplitude reflection coefficients have weak dependence on the polarization direction when the incident angle is small. Hence, the incident beam with a random polarization direction yields $I_o^R(x,y) \simeq I_e^R(x,y)$.

5.1.3 Reconstruction of the phase: normal algorithm

The phase aberration $\delta(x,y)$ is reconstructed from the four interferograms, using the following approximation equations:

$$V_o(x,y) = \frac{I_o^T(x,y)}{I_o^R(x,y)} \left(\frac{R_{\text{out}}^o}{T_{\text{out}}^o}\right)^2 (1+\rho_o f)^2 - 1 - \tau_o^2 f^2 \quad (5.21)$$

$$V_e(x,y) = \frac{I_e^T(x,y)}{I_e^R(x,y)} \left(\frac{R_{\text{out}}^e}{T_{\text{out}}^e}\right)^2 (1+\rho_e f)^2 - 1 - \tau_e^2 f^2 \quad (5.22)$$

$$\delta_{rec}(x,y) = \arctan\left(\frac{V_o(x,y)\tau_e\cos\phi_e - V_e(x,y)\tau_o\cos\phi_o}{V_e(x,y)\tau_o\sin\phi_o - V_o(x,y)\tau_e\sin\phi_e}\right), \quad (5.23)$$

where the function f and the values of the transmission/reflection properties are assumed to be known in advance with analytical calculations (Fig. 5.3) and measurements, respectively, and $|\alpha(x,y)|$ is assumed to be much smaller than 1 and hence is ignored. The normalized interferograms are denoted as $V_o(x,y)$ and $V_e(x,y)$. The reconstructed phase is denoted as $\delta_{rec}(x,y)$. The arctan function is practically implemented by the atan2 function, which does not cause ambiguity in the range of $|\delta| < \pi$ (P-V $< 1\lambda_C$). On the basis of the relation $I_o^R(x,y) \simeq I_e^R(x,y)$, I_o^R and I_e^R are replaced with I_{ave}^R given by,

$$I_{ave}^R(x,y) = \frac{I_o^R(x,y) + I_e^R(x,y)}{2}. \quad (5.24)$$

When I_o^R and I_e^R contain random noise, the substitution of I_{ave}^R for them reduces the noise.

5.1.4 Reconstruction of the phase: iterative algorithm

The normal algorithm presented in the previous section uses the analytical f (Fig. 5.3) to represent the amplitude of the reference beam. However, the amplitude is attenuated in the real PDI when large phase aberrations degrade the Strehl ratio at the focal plane. The iterative algorithm presented in this section takes into account this effect.

First, the phase aberration δ_{rec} is reconstructed with Eqs. 5.21-5.23. Then, the Strehl ratio SR is estimated using the standard deviation s [rad] of the δ_{rec}:

$$SR = \exp\bigl(-s^2\bigr). \tag{5.25}$$

Then, the terms f in Eqs. 5.21-5.23 are replaced with the attenuated f_{SR}:

$$f_{SR} = \sqrt{SR}\ f. \tag{5.26}$$

Then, δ_{rec} is reconstructed again.

Iterating the above procedure yields a more accurate reconstruction of a large phase aberration δ ($\sim \pi$ rad).

5.1.5 Calculation cost

The normal algorithm in Section 5.1.3 calculates the phase aberration $\delta_{rec}(x, y)$ at each subaperture (x, y). Thus, the calculation cost is proportional to the total number of the effective subapertures N_{SA}. In other words, the cost is $\mathcal{O}(N_{SA})$.

The iterative algorithm in Section 5.1.4 calculates $\delta_{rec}(x, y)$ and its standard deviation in each iteration step. The calculation cost of each step is $\mathcal{O}(N_{SA})$. The steps are iterated, but the cost is still proportional to N_{SA}. Thus, the total cost is $\mathcal{O}(N_{SA})$.

5.2 Extension of the principle

The ideal values for the phase differences (θ_o and θ_e) in the b-PDI for the two axes are $\pm 2\pi/3$ rad (section 5.1.1). However the phase differences in practice are not limited to these two values. Therefore it is not necessary for n to satisfy Eq. 5.1. With this less stringent condition, it is easier to find materials suitable for the b-PDI design. In this section we describe the values of θ_o and θ_e with two candidate materials, which are available realistically, and redefine the equations for the phase reconstruction.

5.2.1 Material

We list combinations of materials along with their phase differences in Table 5.2. Combination (A) is preferable because of smaller effect of defocus with smaller d.

Table 5.2: Combinations of materials. The values of refractive indices n_o, n_e and n are defined at $\lambda_C = 800$ nm

	(A)	(B)
Birefringent crystal	TiO_2 (Rutile)	YVO_4
n_o	2.52 [54]	1.97 [55]
n_e	2.79 [54]	2.19 [55]
Non-birefringent material	Nb_2O_5	Ta_2O_5
n	2.28 [56]	2.10 [56]
d	$1.32\lambda_C$	$2.90\lambda_C$
θ_o	-0.634π rad	0.754π rad
θ_e	-1.34π rad	-0.522π rad

5.2.2 Formulation

Some of the parameters are redefined as, using Eq. 5.13 and Eq. 5.14:

$$\tau'_o \exp\bigl(i\phi'_o\bigr) = \frac{T^o_{\text{in}}}{T^o_{\text{out}}} \exp(i\theta_o) - 1 \tag{5.27}$$

$$\tau'_e \exp\bigl(i\phi'_e\bigr) = \frac{T^e_{\text{in}}}{T^e_{\text{out}}} \exp(i\theta_e) - 1. \tag{5.28}$$

Equations 5.21–5.23 are modified as follows.

$$V_o(x,y) = \frac{I^T_o(x,y)}{I^R_o(x,y)} \left(\frac{R^o_{\text{out}}}{T^o_{\text{out}}}\right)^2 (1+\rho_o f)^2 - 1 - \tau'^2_o f^2 \tag{5.29}$$

$$V_e(x,y) = \frac{I^T_e(x,y)}{I^R_e(x,y)} \left(\frac{R^e_{\text{out}}}{T^e_{\text{out}}}\right)^2 (1+\rho_e f)^2 - 1 - \tau'^2_e f^2 \tag{5.30}$$

$$\delta_{\text{rec}}(x,y) = \arctan\left(\frac{V_o(x,y)\tau'_e \cos\phi'_e - V_e(x,y)\tau'_o \cos\phi'_o}{V_e(x,y)\tau'_o \sin\phi'_o - V_o(x,y)\tau'_e \sin\phi'_e}\right) \tag{5.31}$$

5.2.3 Effect of the incident angle

We consider the effect of the incident angle of 10 deg.

First, we consider the lateral displacement of the reference beam to the test beam. Fig. 5.4 shows the schematic view of the relation between incident rays and the BPBS. Incident rays 1 and 2 are polarized in the same direction. Ray 1 enters TiO_2 with the incident angle of 10 deg. Its refraction angle is ϕ_1. By the time it crosses the surface at the depth d, it has laterally displaced by Δ_1 from the incident point. Similarly, Ray 2 enters the pinhole (Nb_2O_5) with the incident angle of 10 deg and its refraction angle is ϕ_2. By the time it crosses the bottom of the pinhole at the depth of d, it has laterally displaced by Δ_2 from the incident point. Because Δ_1 is not equal to Δ_2, the reference beam laterally displaces to the test beam. Calculations using the values in Table 5.2(A) yield the displacement $|\Delta_1 - \Delta_2|$ to be of the order of 10 nm. This value is much smaller than the size of the pinhole, whose diameter is of the order of 10 μm. Thus, the displacement is negligible.

Second, we consider the changes in θ_o and θ_e. When the incident angle is 0 deg, they are given by Eqs. 5.2 and 5.3, respectively. When the incident angle is 10 deg, they are, respectively,

$$\theta_o = \frac{2\pi}{\lambda_C}\left[\left(\frac{nd}{\cos\phi_2} - \frac{n_o d}{\cos\phi_1}\right) + (\Delta_1 - \Delta_2) n_o \sin\phi_1\right] \tag{5.32}$$

$$\theta_e = \frac{2\pi}{\lambda_C}\left[\left(\frac{nd}{\cos\phi_2} - \frac{n_e d}{\cos\phi_1}\right) + (\Delta_1 - \Delta_2) n_e \sin\phi_1\right]. \tag{5.33}$$

Calculations using the values in Table 5.2(A) yield the result that the change in the incident angle by 10 deg deviates θ_o and θ_e by 0.3% and 0.2%, respectively. Thus, the changes in θ_o and θ_e are negligible.

5.3 Summary of this chapter

This chapter has proposed the new phase sensor, the b-PDI, and determined its reconstruction equations and algorithms. The calculation cost is $\mathcal{O}(N_{SA})$. The size of the readout region is N_{1d} pixels in the vertical direction. The cost and size are the smallest level among existing phase sensors and are advantageous for high-speed measurement. The reconstruction equation is applicable without ambiguity in the P-V $< 1\lambda_C$ range; it is two times larger than other phase sensors. The b-PDI is expected to have high efficiency, which is examined by simulations in the next chapter.

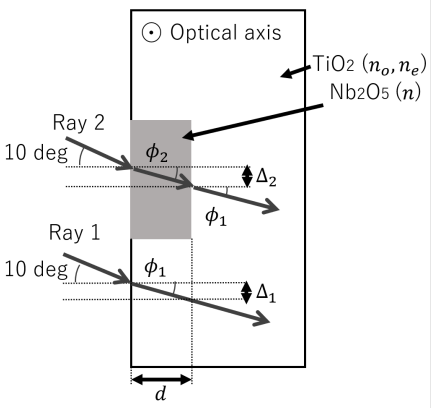


Figure 5.4: Sectional view of the BPBS with incident rays. Reprinted with permission from [51] ©Optica Publishing Group.

Chapter 6

Simulations and Optimal Design

This chapter examines the efficiency and other error sources of the b-PDI. In addition, this chapter discusses the optimal pinhole size. The simulation in this chapter uses the normal algorithm (Section 5.1.3).

This chapter is based on Sections 4 and 5 of Tsukui et al. [51], published in Applied Optics. Note that the simulations after Section 6.1.3 of this thesis adopt the pinhole radius $R_p = 15$ μm, whereas Tsukui et al. [51] adopted $R_p = 12$ μm. Additionally, the discussion on the sensitivity factor β_p in Section 6.1.2 is revised.

6.1 Simulations

We ran numerical simulations to determine the optimal R_p and to estimate the expected measurement errors. In addition, we evaluated the Strehl ratio and contrast achieved with the ExAO that employs the b-PDI. we assumed combination (A) in Table 5.2 as the materials of the BPBS and used the custom-developed software utilizing NumPy and SciPy. The four components of measurement error were considered, i.e., those originating from

- approximation in Eqs. 5.29–5.31: σ_{rec}
- noise: $\sigma_{stat} = \sqrt{\sigma_{pn}^2 + \sigma_{rn}^2}$
  - photon noise: σ_{pn}
  - readout noise: σ_{rn}

- chromatic effect: σ_{chrom}.

The total measurement error, σ_{WFS}, is described as

$$\sigma_{WFS} = \sqrt{\sigma_{rec}^2 + \sigma_{stat}^2 + \sigma_{chrom}^2}. \tag{6.1}$$

6.1.1 Setup

We express the incident wavefront aberration $W(x, y)$ with the Zernike polynomials $Z_n^m(x, y)$. The aberration is converted into the phase aberration $\delta(x, y)$:

$$\delta(x, y) = \frac{2\pi}{\lambda_C} W(x, y) \tag{6.2}$$

$$= \frac{2\pi}{\lambda_C} \sum_{n=0}^{N} \sum_{m=-n}^{n} a_n^m Z_n^m(x, y), \tag{6.3}$$

where a_n^m is the Zernike coefficients. With this, the P-V value of each mode normalized with λ_C is $2a_n^m/\lambda_C$ as in Eq. 6.3. Setting the amplitude to 1 across the input pupil, we express the complex amplitude by

$$E(x, y) = \exp\{i\delta(x, y)\}. \tag{6.4}$$

The complex amplitude is converted into the four interferograms with a numerical model of the b-PDI. The model is based on Fraunhofer diffraction and includes the following series of calculations (see Fig. 6.1 for summary). First, the array of the complex amplitude (240-pix diameter) is padded with zeros [57]. Second, the padded complex amplitude is converted into a complex amplitude at the focal plane with fast Fourier transformation (FFT). The resolution of the focal plane is sufficiently good (1 pix = 0.5 μm) due to the zero padding on the pupil plane. Third, the complex amplitude is multiplied by the transmission properties or the reflection properties of the BPBS given in Table 6.1. This process simulates the phase shifts and the transmission/reflection of the BPBS. The four maps of the complex amplitude corresponding to e/o and transmission/reflection are generated in this process. Fourth, and lastly they are converted into four interferograms (240-pix diameter) with inverse fast Fourier transformation (IFFT) and conversion to the intensity. The wavefront aberration of the reference beam, which is related to the fact that the pinhole size is finite, is taken into account in this calculation. Each interferogram is binned to a resolution of

24×24. This means that the aperture is resolved into 24 subapertures in diameter, i.e., $N_{1d} = 24$. The resultant set of the four interferograms are $I_o^T(x,y)$, $I_e^T(x,y)$, $I_o^R(x,y)$, and $I_e^R(x,y)$. The phase is reconstructed from them with Eqs. 5.29–5.31. The reconstructed phase aberration $\delta_{\rm rec}(x,y)$ (24-pix diameter) is converted into the wavefront aberration $W_{\rm rec}(x,y)$ (24-pix diameter), and then the measurement error is obtained by comparing $W_{\rm rec}(x,y)$ and $W(x,y)$ (binned to a resolution of 24×24).

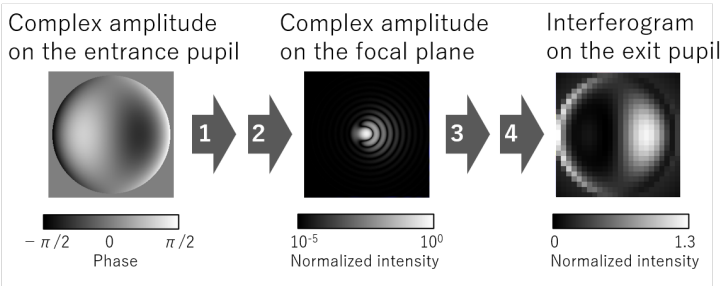


Figure 6.1: Numerical model of the b-PDI. 1: zero padding, 2: FFT, 3: transmission (including the phase shift) or reflection, 4: IFFT and conversion to the itensity. Reprinted with permission from [51] ©Optica Publishing Group.

Table 6.1: Transmission and reflection properties at the measurement wavelength λ_C = 800 nm. The values of T/R are the Fresnel coefficients calculated from the refractive indices in Table 5.2 (A) and the incident angle of 10 deg. The values of θ_o and θ_e are same as those in Table 5.2 (A)

Polarization:	Parallel to the x-axis (o)	Parallel to the y-axis (e)
Transmission		
Inside [1]	$T_{\rm in}^o \ \exp(i\theta_o) = 0.922\exp(i\theta_o)$	$T_{\rm in}^e \ \exp(i\theta_e) = 0.913\exp(i\theta_e)$
Outside [2]	$T_{\rm out}^o = 0.904$	$T_{\rm out}^e = 0.879$
Reflection		
Inside [1]	$R_{\rm in}^o = 0.385$	$R_{\rm in}^e = -0.395$
Outside [2]	$R_{\rm out}^o = 0.427$	$R_{\rm out}^e = -0.477$

[1] Inside of the pinhole
[2] Outside of the pinhole

6.1.2 Effect of pinhole size

The measurement error depends on R_p. We sought the optimal R_p to minimize the measurement error with simulations. In the simulations, We assumed monochromatic beams with a wavelength of $\lambda_C = 800$ nm and an Airy disk with a radius of $R_{Airy} = 30$ μm. Under these condition, I estimated the measurement errors σ_{rec} and $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$. The former, σ_{rec}, is a measurement error in a special case in which no noise is present. The latter, $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$, is a measurement error in a general case in which the four interferograms contain photon noise, which was expressed with random numbers following a Poisson distribution. The number of the incident photons (N_{ph}) in one subaperture in the simulations was 100, a half of which (50 photons) contributed to I_o^T or I_o^R, and the other half of which contributed to I_e^T or I_e^R. Ideal detectors with 100-% QE (quantum efficiency) and no readout noise were assumed.

Figs. 6.2 and 6.3 show the resultant measurement errors with five R_p. The first part of the measurement errors, σ_{rec}, was an increasing function of R_p when the incident aberration was larger than $0.2\lambda_C$. In contrast, the second part of the measurement errors, $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$, was minimized at a certain value of R_p. We chose 15 μm for R_p to minimize the error for the incident aberrations with P-V $\leq 0.5\lambda_C$ as shown in Fig. 6.3. The value corresponds to 50% of the radius of the Airy disk (R_{Airy}). We adopted this value in the following simulations and discussion.

With this value of R_p, the measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ was 7.7×10^{-2} rad when the incident aberration was $0.01\lambda_C$ and the number of the incident photons was 100. Considering $\sigma_{rec} = 3.0 \times 10^{-3}$ rad, σ_{pn} is estimated to be 7.7×10^{-2} rad. This indicates that the coefficient of the error propagation C_p (Eq. 4.4) is 0.77. The value is $\simeq 1.1$ times larger than C_p of the ZWFS ($\sqrt{0.5}$, Table 4.1). This means that the b-PDI is $\simeq 1.1$ times more sensitive to photon noise than the ZWFS. Thus, the sensitivity factor β_p of the b-PDI is estimated to be $\simeq 1.1$.

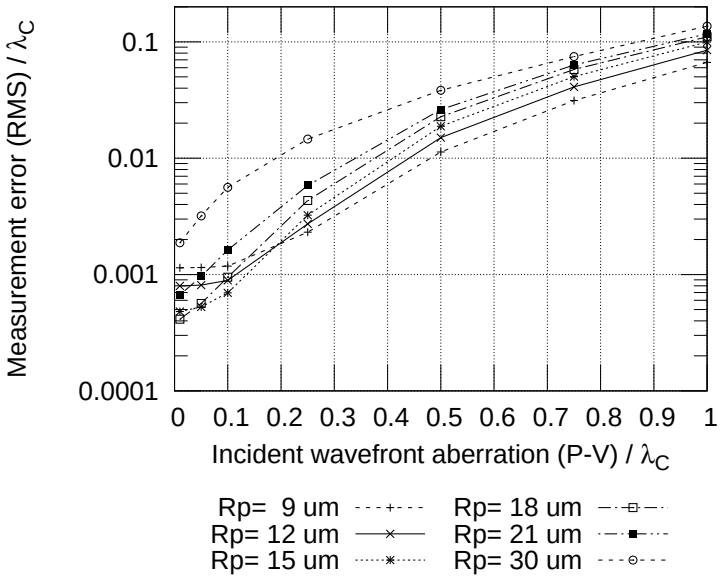


Figure 6.2: Measurement error σ_{rec}, plotted against the P-V value of the incident wavefront aberration expressed with the Zernile polynomial with $n = 3, m = 1$ (Coma X). Reprinted with permission from [51] ©Optica Publishing Group.

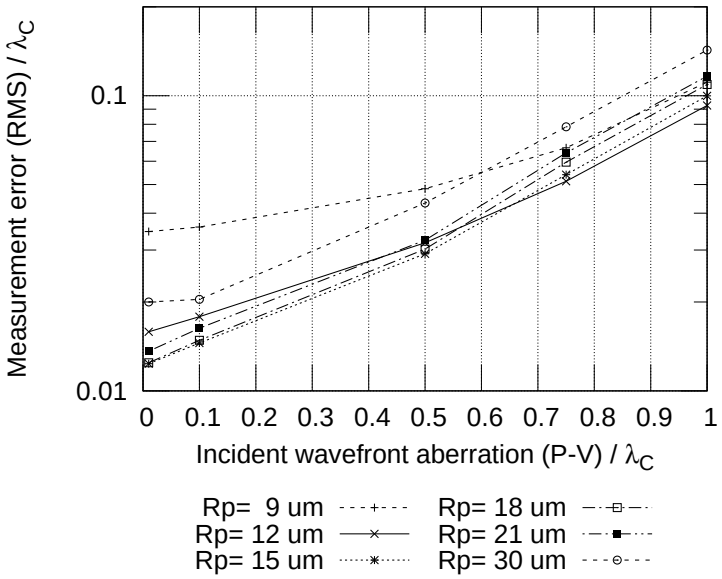


Figure 6.3: Measurement error $\sqrt{\sigma_{\mathrm{rec}}^2 + \sigma_{\mathrm{pn}}^2}$ (median of 1000 trials), plotted in the same way as in Fig. 6.2. Reprinted with permission from [51] ©Optica Publishing Group.

6.1.3 Effect of Misalignment

We evaluated the effect of misalignment of the BPBS. We considered the cases with the decenter values of 2, 4, and 6 μm or the defocus of 300 μm for the misalignment and estimated the measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$. Note that the decenter is a misalignment perpendicular to the axis of L0 (Fig. 5.2), and the defocus is a misalignment along the axis. In the cases of decenter, the reconstructed wavefront $W_{rec}(x, y)$ contained a false tilt aberration, while in the case of defocus, $W_{rec}(x, y)$ contained a false defocus aberration. We subtracted these aberrations and then evaluated the measurement error, assuming the same wavelength and the same number of photons as in the previous simulations. The incident wavefront aberration was expressed with the Zernike polynomial of coma ($n = 3,\ m = 1$), which is orthogonal to the tilt ($n = 1,\ m = 1$) or defocus ($n = 2,\ m = 0$).

Fig.6.4 shows the derived measurement errors for the decentered positions for the BPBS. The BPBS decenter by 2, 4, and 6 μm increased the error by 13, 38, and 62 %, respectively. Thus, to maintain the measurement accuracy, the decenter must be kept smaller than a few microns. To

achieve this, the temperature of this system must be stabilized to avoid lateral misalignment due to thermal expansion.

Fig.6.5 shows the measurement error for the defocused position for the BPBS. The BPBS defocus by 300 μm increased the error by 13 %. Though the amount of defocus is much larger than that of the decenters examined above, the defocus has a smaller effect on the measurement accuracy. Thus, an acceptable amount of defocus is about 100 times larger than that of decenter.

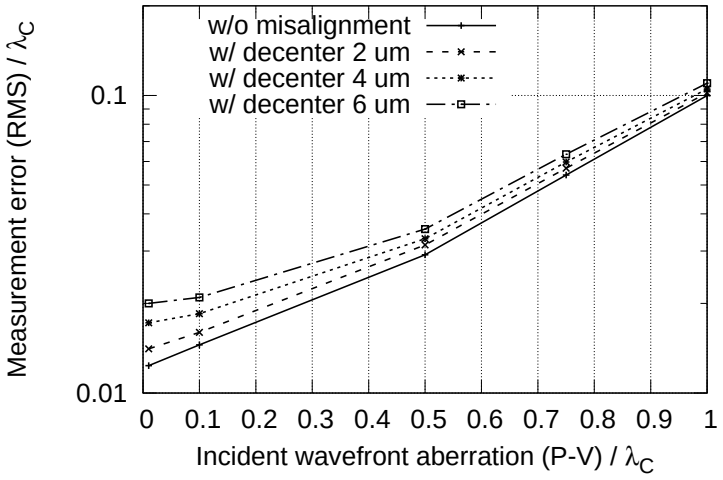


Figure 6.4: Measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ (median of 1000 trials) with the BPBS decentered for 3 positions, plotted in the same way as in Fig. 6.2.

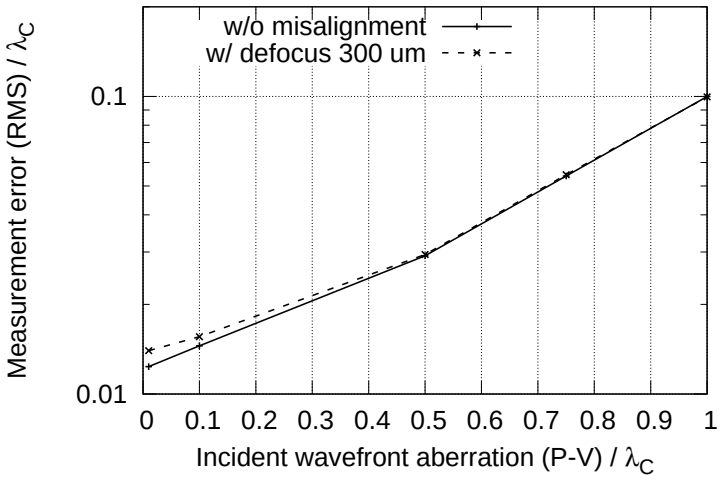


Figure 6.5: Measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ (median of 1000 trials) with the defocused BPBS, plotted in the same way as in Fig. 6.2.

6.1.4 Chromatic Effect

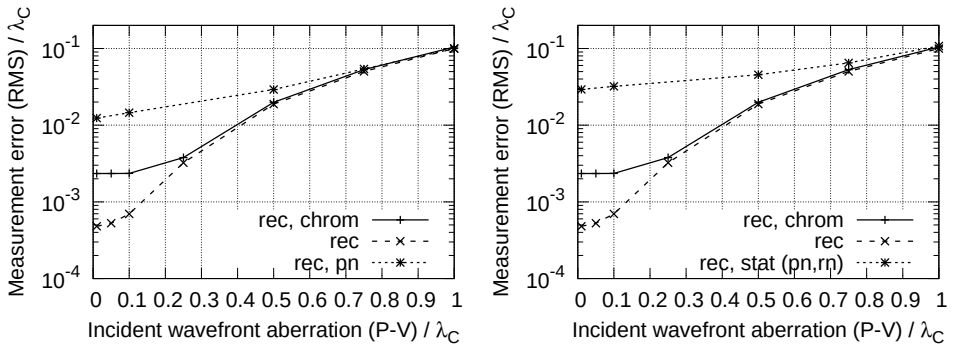
The incident beam in real observation has a finite bandwidth to secure the number of photons under the low photon flux from guide stars. We evaluated the chromatic effect caused by the bandwidth $\Delta\lambda$ of the incident beam. Assuming the central wavelength $\lambda_C = 800$ nm and a flat spectrum over $\Delta\lambda = 200$ nm, We executed the following simulation. First, five sets of interferograms were generated independently with the numerical model of the b-PDI and wavefront aberration $W(x, y)$. The five sets of interferograms were monochromatic and corresponded to wavelengths of 700, 750, 800, 850, and 900 nm. In this process, We considered chromatic effects such as difference in the Airy radius and transmission/reflection properties. Second, the interferograms of the different wavelengths were averaged with equal weighting to generate a set of polychromatic interferograms: $I_o^T(x, y)$, $I_e^T(x, y)$, $I_o^R(x, y)$ and $I_e^R(x, y)$. Then, the phase was reconstructed from the polychromatic interferograms. Finally, the phase aberration was converted to the wavefront aberration W_{rec} as

$$W_{\mathrm{rec}}(x, y) = \frac{\lambda_C}{2\pi}\delta_{\mathrm{rec}}(x, y) \tag{6.5}$$

and the measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{chrom}^2}$ was calculated.

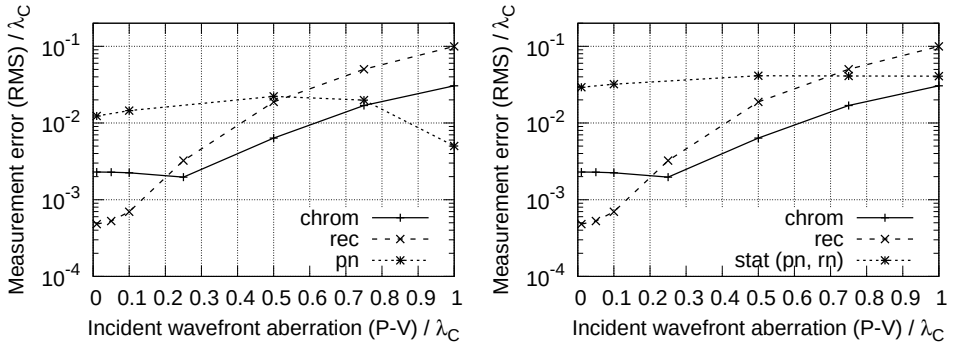
The measurement error is compared in Fig. 6.6 with the errors σ_{rec},

$\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ and $\sqrt{\sigma_{rec}^2 + \sigma_{stat}^2}$. The error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ was evaluated with the photon noise, the QE of an ideal detector (= 1.0), and $N_{ph} = 100$. The error $\sqrt{\sigma_{rec}^2 + \sigma_{stat}^2}$ was evaluated with the photon noise, the readout noise (1.6 electrons RMS), the QE of a real detector (= 0.5) [36], and $N_{ph} = 100$. These errors are decomposed into each component in Fig. 6.7. When $N_{ph} = 100$, $\sigma_{chrom}^2 \ll \sigma_{rec}^2 + \sigma_{pn}^2$ and $\sigma_{chrom}^2 \ll \sigma_{rec}^2 + \sigma_{stat}^2$. Therefore, σ_{chrom} is negligible in this case.



(a) **(b)**

Figure 6.6: (a)Measurement errors σ_{rec} (dashed line), $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ (median of 1000 trials, dotted line), and $\sqrt{\sigma_{rec}^2 + \sigma_{chrom}^2}$ (solid line), plotted in the same way as in Fig. 6.2. (b)Measurement errors σ_{rec} (dashed line), $\sqrt{\sigma_{rec}^2 + \sigma_{stat}^2}$ (median of 1000 trials, dotted line), and $\sqrt{\sigma_{rec}^2 + \sigma_{chrom}^2}$ (solid line), plotted in the same way as in Fig. 6.2.



(a) **(b)**

Figure 6.7: (a)Comparison of each component of the measurement errors σ_{rec} (dashed line), σ_{pn} (dotted line), and σ_{chrom} (solid line), plotted as in Fig.6.6. (b)Comparison of each component of the measurement errors σ_{rec} (dashed line), σ_{stat} (dotted line), and σ_{chrom} (solid line), plotted as in Fig.6.6.

6.1.5 Simulation of ExAO

We simulated a closed-loop ExAO on a 3.8-m telescope that contains the b-PDI to evaluate the achievable Strehl ratio and contrast. We supposed an ExAO consisting of two AO systems connected in series: a Woofer AO and a Tweeter AO. The Woofer AO contained a SHWFS with 8×8 subapertures and corrected low-order wavefront aberrations. The Tweeter AO contained the b-PDI with 24×24 subapertures and corrected the residual wavefront aberrations of the Woofer AO.

The aberration of the incident wavefront for the Tweeter AO was 1200 nm (P-V), under an assumption that the wavefront was corrected in advance with the Woofer AO. The incident wavefront was generated with power-spectrum-based simulations by Yamamoto. Details of the wavefront and the Woofer AO are listed in Table 6.2.

We used a Tweeter AO model, which consisted of a wavefront sensor model and a deformable mirror (DM) model. As for the former (wavefront sensor model), we used the numerical model of the b-PDI described in Section 6.1.1 and considered a single wavelength of $\lambda_C = 800$ nm on the basis of the result in Section 6.1.4. Photon noise and/or readout noise were added to the four interferograms. Some peaks and valleys in the incident wavefront exceeded the range of λ_C. However, such regions were smaller than a subaperture of the wavefront sensor. Therefore, we did not execute phase-unwrapping in the measurement with the b-PDI. As for the latter (DM model), the figure of the continuous surface of a DM was modeled using spline interpolation of a stroke map, whose resolution was 24×24. Detailed conditions are summarized in Table 6.3.

The i-th AO loop included the following operations. First, the incident wavefront $W_{Inc240,i}$ (240-pix diameter) was corrected by the DM model with the surface $W_{DM240,i-1}$ (240-pix diameter), which had been determined in the previous $(i-1)$-th loop. The corrected wavefront $W_{Cor240,i}$ (240-pix diameter) was expressed as:

$$W_{Cor240,i} = W_{Inc240,i} - W_{DM240,i-1}. \tag{6.6}$$

Then, the corrected wavefront was measured by the b-PDI as $W_{rec24,i}$ (24-pix diameter). Then, the stroke map of the DM actuators $U_{24,i}$ (24-pix diameter) was updated with the following integral controller:

$$U_{24,i} = U_{24,i-1} + K_I W_{rec24,i}, \tag{6.7}$$

where K_I (a integral gain) was set to 0.3. Finally, the updated DM surface $W_{DM240,i}$ (240-pix diameter) was determined using spline interpolation of the stroke map $U_{24,i}$.

We assumed a high loop-speed of 6.5 kHz and considered the lateral drift of the incident wavefront $W_{inc240,i}$ due to assumed 10-m/s wind. Atmospheric chromaticity was ignored to evaluate the effect of the errors in wavefront sensing, i.e., the WFS noise error, temporal error and aliasing error.

The loop was successfully closed on the incident wavefront with 15 iterations. The aberration of the corrected wavefront was $\lesssim 80$ nm RMS, which corresponds to an H-band Strehl ratio of 90%. The corrected wavefront at the central region with a diameter of 3.5 m was extracted with a Lyot stop and was converted into an H-band ($\lambda_i = 1650$ nm) PSF with FFT. Fig. 6.8 shows the resultant wavefront maps and PSFs. Figs. 6.9 and 6.10 show the averaged contrast as a function of the angular separation. Contrasts of $\sim 10^{-5}$ with an angular separation of 0.2 arcseconds was achieved.

Table 6.2: Details of the incident wavefront and the Woofer AO

Fried parameter r_0 (wavelength: 600 nm)	0.15 m
Wind speed v	10 m/s
Woofer AO	
Number of subapertures/actuators	8×8
Wavefront sensor type	SHWFS
Frame rate	1 kHz
Wavefront error before the Tweeter AO	1200 nm P-V, 150 nm RMS

Table 6.3: Details of the Tweeter AO model. Real or ideal detectors were assumed. The values of the real detectors' QE and readout noise are based on Hamamatsu ORCA-Flash 4.0v2 sCMOS camera [36].

Telescope diameter	3.8 m
Throughput of the optics before Tweeter WFS	0.24
Magnitude of the guide star	4.8 mag (Ic-band)
Integral controller's gain K_I	0.3
Wavefront sensor model (b-PDI)	
Sensing wavelength λ_C	800 nm
Radius of the Airy disk R_{Airy}	30 μm
Radius of the pinhole R_p	15 μm
Number of subapertures	24×24
Frame rate	6.5 kHz
Number of photons per subaperture	100 (50 for each polarization)
Quantum efficiency (QE) of the detectors	1.0 (ideal)/ 0.5 (real)
Readout noise of the detectors	0 e^- (ideal)/ 1.6 e^- RMS (real)
Deformable mirror model	
Number of actuators	24×24
Maximum stroke	1200 nm

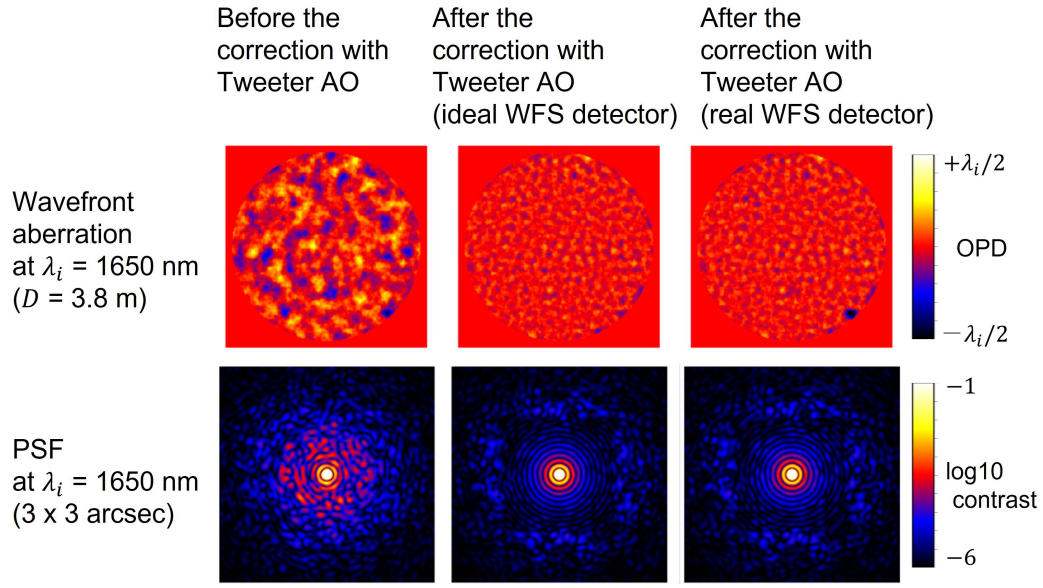


Figure 6.8: Wavefront maps with a diameter of 3.8 m and PSFs with a 3×3 arcsec region. Left column: no Tweeter correction is applied. Center column: the Tweeter correction with the ideal WFS detectors is applied. Right column: the Tweeter correction with the real WFS detectors is applied.

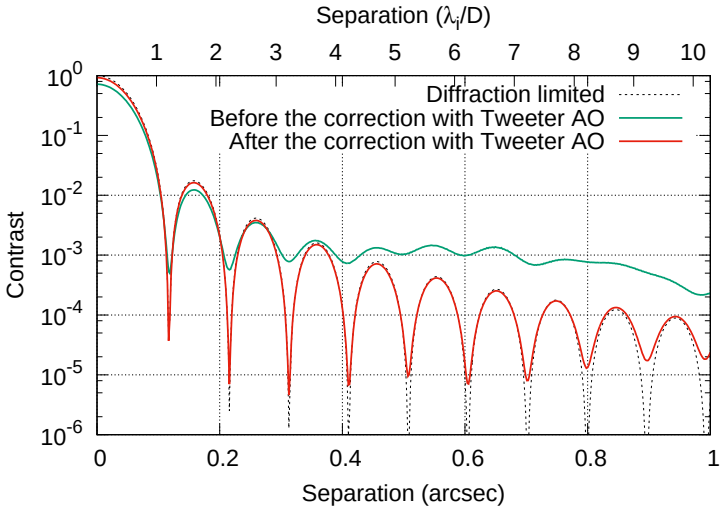


Figure 6.9: Averaged contrast of the PSF, which is the azimuthally averaged intensity normalized by the peak intensity of the diffraction-limited PSF, before and after the correction with the Tweeter AO (with the ideal WFS detectors), as a function of the angular separation from the peak. The profile of the diffraction-limited PSF (dotted line) is shown for reference.

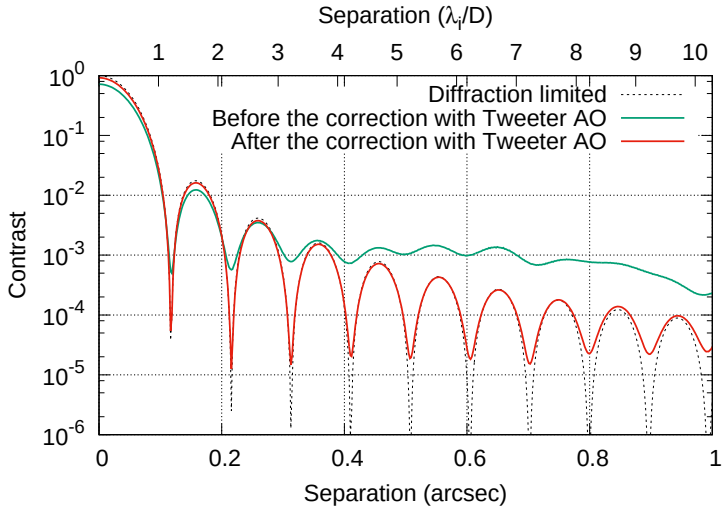


Figure 6.10: Averaged contrast of the PSF, which is the azimuthally averaged intensity normalized by the peak intensity of the diffraction-limited PSF, before and after the correction with the Tweeter AO (with the real WFS detectors), as a function of the angular separation from the peak. The profile of the diffraction-limited PSF (dotted line) is shown for reference.

6.2 Discussion on the optimal design

We examine the relation between the pinhole size R_p and the measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ with numerical simulations. In the simulations, we considered the change in the amount of the photon noise as a function of R_p. As a result, a certain value of R_p yielded the minimum measurement error, balancing the following three factors.

First, aberration of the reference beam is induced at the pinhole. In the wavefront reconstruction with the PDI, it is assumed that the reference beam is not aberrated. This assumption holds only when the pinhole radius is infinitely small. In reality, however, an actual pinhole has a finite radius. Thus, it extracts components with non-zero spatial frequencies in addition to that with zero spatial frequency from the incident beam. As a result, the reference beam is aberrated and deviates from the assumed non-aberrated beam. This contributes to the measurement error. Larger the pinhole is, more components with higher spatial frequencies contaminate the reference beam, thus resulting in a larger aberration of the reference beam. Consequently, the measurement error σ_{rec} is an increasing function of R_p.

Second, the contrast of the interferograms varies, depending on the conditions. The contrast is maximum (= 1) when $R_p \sim 0.5F\lambda_C$, where the energy of the beams inside the pinhole is equal to that outside. When the interferograms contain photon noise, their S/N ratio is an increasing function of the contrast, whereas the measurement error σ_{pn} is a decreasing function of the S/N ratio. Hence, at $R_p \sim 0.5F\lambda_C$, the S/N ratio is maximized and the measurement error σ_{noise} is minimized.

The third factor is attributed to Eq. 5.24. Smaller the R_p is, a wider region on the outside of the pinhole contributes to the interferograms. The difference in reflectance between the two polarization directions on the outside of the pinhole is larger than that inside. Thus, a smaller pinhole increases the deviation of I_o^R from I_e^R, hence making the assumption of Eq. 5.24 less appropriate and resulting in a larger measurement error σ_{rec}. This effect is apparent in the region with the incident aberration of smaller than $0.2\lambda_C$ in Figure 6.2.

In practice, the measurement error $\sqrt{\sigma_{rec}^2 + \sigma_{pn}^2}$ is minimized at a certain value of R_p as a balance between the first and second factors only, because the contribution of the third factor is negligible being much smaller than the second factor. The simulations finds that R_p of $\simeq 15$ μm ($\simeq 0.5R_{Airy}$) is acceptable with a reasonably small measurement error and that the measurement error rapidly increases for smaller R_p when $R_p \leq 12$ μm. Note that

R_p needs to be fairly precisely controlled when the pinhole is manufactured.

I15 examined the relation between the measurement error of the interferograms and that of the wavefront by analyzing error propagation and pointed out that R_p of $0.4 - 0.6R_{Airy}$ would suppress the error propagation. Though they did not consider the varying amount of the photon noise depending on R_p, their result is consistent with ours.

6.3 Summary of this chapter

This chapter has predicted the b-PDI efficiency under the optimum pinhole radius $R_p = 15$ μm. This chapter also has examined other error sources such as chromatic effect. Table 6.4 compares the b-PDI's simulated performance and target performance (Table 5.1). Table 6.5 compares the simulated b-PDI with other phase sensors (Table 4.1).

Table 6.4: Target and simulated performance of the b-PDI. β_p: sensitivity factor [27]. N_{1d}: 1-dimensional number of the effective subapertures. λ_C: center of the sensing wavelength band. $\Delta\lambda$: width of the sensing wavelength band.

	Target	Simulation
Type	Phase sensor	Phase sensor
β_p	≤ 1.4	$\simeq 1.1$
Sampling rate	6.5 kHz [i]	6.5 kHz
Calculation cost	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA})$
Number of detector pixels (vertical)	N_{1d}	N_{1d}
N_{1d}	≥ 24 [i]	24
Dynamic range	$1\lambda_C$	$1\lambda_C$ [ii]
λ_C	800 nm [i]	800 nm

[i] Assuming the SEICA instrument [17].
[ii] The range in which the reconstruction equation is applicable without ambiguity.

Table 6.5: Comparison of existing WFSs. C_p: coefficient of error propagation from photon noise. β_p: sensitivity factor [27]. λ_C: center of the sensing wavelength band. N_{SA}: number of all the effective subapertures. N_{1d}: 1-dimensional number of the effective subapertures.

	ZWFS	FPyWFS	b-PDI (simulated)
Type	Phase	Phase	Phase
C_p	$\sqrt{0.5}$ [i]	N/A	0.77
β_p	1.0	$\sqrt{2}$	$\simeq$1.1
Calculation cost	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA}) - \mathcal{O}(N_{SA}^2)$	$\mathcal{O}(N_{SA})$
Number of detector pixels (vertical)	N_{1d}	$> 2N_{1d}$	N_{1d}
Dynamic range	$< 0.5\lambda_C$ [ii]	$\lesssim 0.5\lambda_C$ [ii]	$1\lambda_C$ [ii]

[i] From N'Diaye et al. 2013 [48].

[ii] The range in which the reconstruction equation is applicable without ambiguity.

Chapter 7

Manufacturing

This chapter describes the manufacturing of the BPBS and b-PDI. The BPBS was manufactured under the design in Chapters 5 and 6. This chapter is based on Section 3 of Tsukui et al. [58], to be published in Optics Continuum.

7.1 BPBS

We manufactured the BPBS. Based on Tsukui et al. [51], the pinhole was designed with a depth of $d = 1.32\lambda_C = 1.06$ μm and a radius of $R_p = 15$ μm. This R_p is 0.5 times the Airy disk radius at an F-number of 31 and a central wavelength of $\lambda_C = 800$ nm. We made another pinhole of radius 200 μm filled with Nb_2O_5 to measure the parameters T and R. Hereafter, the pinholes of radii 15 μm and 200 μm are called Pinholes A and B, respectively. We manufactured 24 Pinholes A and 24 Pinholes B on the front surface of a 20-mm square, 0.5-mm thick TiO_2 substrate. The back of the substrate is AR-coated; the reflectance is less than 1%. We processed eight substrates to select the best Pinhole A from 192 samples.

Figure 7.1 shows an example of the manufacturing process. In the process, the pinholes are first manufactured on the front surface of the TiO_2 substrate by SF_6-based reactive ion etching (RIE) [59]. The etching mask consists of two layers: photoresist (AZ5214E) and Cr. Then, the mask is removed by wet etching with acetone and nitric-acid-based etchant. Then, Nb_2O_5 is deposited on the front surface by an RF sputtering method. Finally, the front surface is manually polished, leaving the Nb_2O_5 film only inside the pinholes. Most parts of the process were performed in the clean room (Mechano-Microprocess Room) at Tokyo Institute of Technology by

the author and the collaborators. The photographs taken in the process are shown as Fig. 7.2(a) and (b).

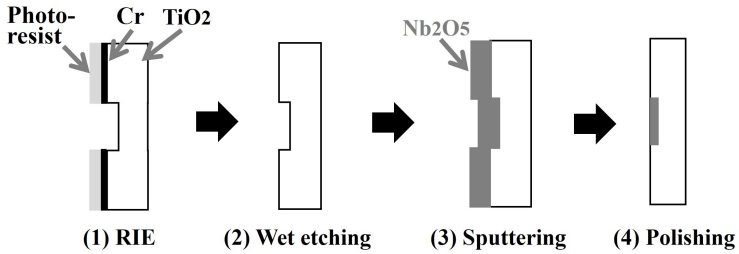


Figure 7.1: Schematic drawing of the manufacturing process. RIE: Reactive Ion Etching. Reprinted with permission from [58] ©Optica Publishing Group.

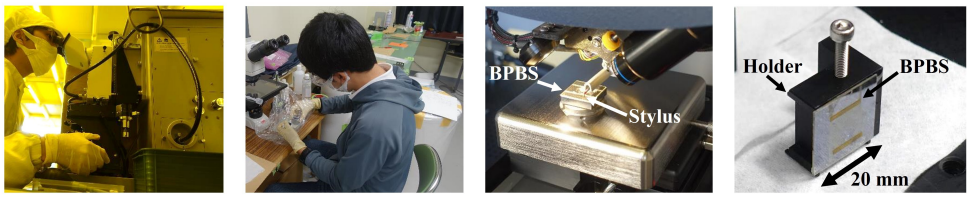


Figure 7.2: (a) Ryo Tsukui operating a mask aligner in the clean room to form the etching mask. (b) Ryo Tsukui conducting an experimental manual polishing. (c) The BPBS set in the stylus profilometer. (d) The manufactured BPBS set on a custom-made holder.

After the process, the shape of the front surface was measured with a stylus profilometer (Bruker DektakXT, Fig. 7.2(c)) at Tokyo Institute of Technology. The thickness t and refractive index n of the Nb_2O_5 film were measured with a microscopic spectrophotometer (JASCO MSV-5200) at Kyoto Prefectural Technology Center. Based on these measurements, We selected several Pinholes A with a shape close to the designed value.

The selected multiple Pinholes A were set into the optical system described in Section 7.2. A tilted flat wavefront was injected into the optical system, and the contrast of the interferograms was evaluated. We selected Pinhole A (#03-40), which produced the best contrast.

Figure 7.3 and Table 7.1 show the shape and phase shifts of Pinhole A (#03-40). The surface of the Nb_2O_5 film is concave by 25 nm relative to the

TiO_2 surface. This concavity causes a phase shift of 0.1π in the reflection path, which is acceptable. The phase shifts in the transmission path are redefined by considering the concavity h as follows:

$$\theta_o = \frac{2\pi}{\lambda_C}\{(n - n_o)t + (1 - n_o)h\}, \tag{7.1}$$

$$\theta_e = \frac{2\pi}{\lambda_C}\{(n - n_e)t + (1 - n_e)h\}. \tag{7.2}$$

The term n, t, and h are measured values. The term n_o and n_e are literature values [54].

Pinhole B (#03-41) adjacent to Pinhole A (#03-40) is also shown in Fig. 7.3 and Table 7.1. The surface of the Nb_2O_5 film is a loose concave surface (P-V $\sim$ 70 nm). The thicknesses of the films in Pinholes A and B are greater than the coherent length of the light used for the measurement of R and T (wavelengths of 800 ± 100 nm). Thus, the films do not cause interference inside them. Therefore, the transmission and reflection coefficients in Pinhole B (#03-41) are uniform and equivalent to those in Pinhole A (#03-40).

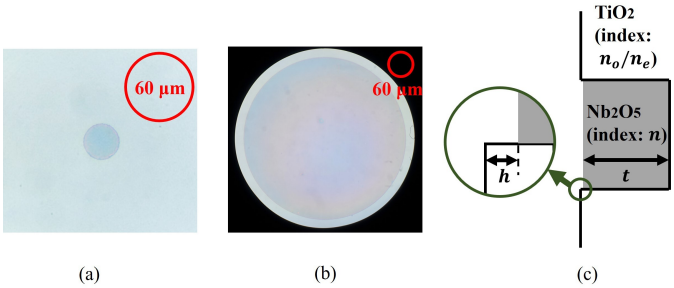


Figure 7.3: (a) Optical microscopic image of Pinhole A (#03-40). (b) Optical microscopic image of Pinhole B (#03-41), in which the black areas are outside the field of view. In (a) and (b), red circles show the size of the Airy disk with a radius of 30 μm (a diameter of 60μm). The pinholes are illuminated with a white LED light source (incident type). (c) Schematic drawing of the cross-section of the manufactured pinhole. Note that the figure is stretched in the direction of the pinhole depth. Reprinted with permission from [58] ©Optica Publishing Group.

We show the optical performance of Pinhole A (#03-40). The optical path difference (OPD) map of the light transmitted through Pinhole A (#03-40) was captured with a transmitted dual-beam interference microscope (MIZOJIRI TD-series) at MIZOJIRI OPTICAL CO., LTD. The wavelength λ_m of the light used in the evaluation was 546 nm. Figure 7.4

Table 7.1: Comparison between the manufactured pinholes and the design. PH: pinhole.

	Designed PH A	Manufactured PH A (#03-40)	Designed PH B	Manufactured PH B (#03-41)
R_p (μm)	15	15	200	200
n	2.28	2.27	2.28	2.27
t (μm)	1.06	0.82	1.06	0.76 (at the center)
h (μm)	0	0.025	0	0.02 to 0.09
θ_o (rad)	-0.634π	-0.61π	–	–
θ_e (rad)	-1.34π	-1.2π	–	–

shows the result. The flatness of the OPD in Pinhole A is 20 nm RMS ($\lambda_m/27 \sim \lambda_C/40$).

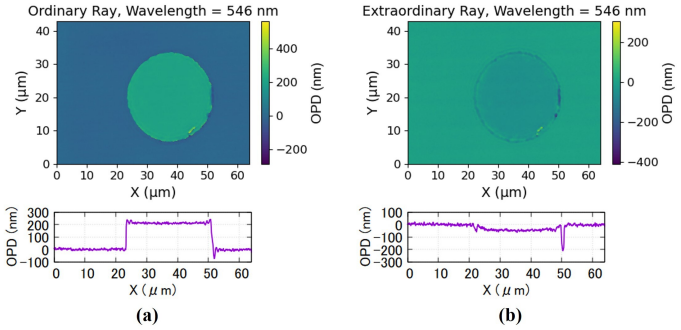


Figure 7.4: OPD maps and their cross-section graphs of Pinhole A (#03-40). (a) Ordinary ray. (b) Extraordinary ray. Reprinted with permission from [58] ©Optica Publishing Group.

7.2 Optical system of the b-PDI

Aiming to mount the b-PDI on the ExAO instrument SEICA [17] of the Seimei Telescope [34], we built an optical system for the b-PDI. The WFS in SEICA requires $\gtrsim 25$ subapertures in the diameter of the telescope aperture because $D/r_0 = 25$, where D is the diameter of the telescope aperture (= 3.8 m) and r_0 is the assumed Fried parameter at the observation site (=

0.15 m). The WFS also requires as few subapertures as possible (i.e. as wide subapertures as possible) to collect adequate photons from stars at each subaperture. Following these requirements, the b-PDI optical system has 492 (26×26) subapertures. The scalability and limiting factor of the subaperture number are discussed in Section 8.2.5.

The system is a modification of the optical system shown in Chapter 5, which requires two detectors. The proposal to use two detectors is unsuitable for the target measurement frequency of 6.5 kHz. This is due to the difficulty of synchronization at the order of 6 kHz with the detectors in hand. To ensure the four interferograms are captured simultaneously, we designed the optical system in which the four interferograms appear on an sCMOS detector (Hamamatsu ORCA-Flash4.0V2). Furthermore, to maximize the frame rate (up to > 6.5 kHz), we minimized the vertical size of the readout region by aligning the four interferograms horizontally on the sCMOS detector.

The layout of the optical system is shown in Fig. 7.5. In the following, the x and y axes are perpendicular to the local axis of the beam and are horizontal and vertical, respectively. The polarization directions of the ordinary ray o and the extraordinary ray e are parallel to the x and y axes, respectively, as shown in Fig. 5.1. The incident beam (6.5-mm diameter) is focused by L0 and L0' to create an Airy disk on the BPBS Pinhole A. The converged beam is F/31, and the Airy disk radius is 30 μm at $\lambda_C = 800$ nm. The BPBS transmits and reflects the beam. The transmitted beam is bent at M0, collimated at L1, and further bent at M1. L1 creates the intermediate pupil IP1. On the other hand, the reflected beam is collimated at L2 and bent at M2. L2 creates the intermediate pupil IP2. The pupils IP1 and IP2 are at the same distance from L3. The beams from IP1 and IP2 pass HWP and SP, which separate the polarization directions o and e. The function of HWP is described in detail later. The separated beams are relayed onto the detector by L3 and L4, creating four interferograms (Fig. 7.6). Each of the interferograms has a diameter of 26 pix (= 169 μm). Under these conditions, the Nyquist frequency is 13 cycles/pupil (c/p). The phase aberrations at > 13 c/p are cut off by a square mask, which is placed in front of the BPBS and limits its clear aperture to 600×600 μm. The BPBS Pinhole A is located at the center of the clear aperture. Figure 7.7 is the photograph of the optical system.

In the optical path, HWP rotates the polarization direction. The polarization direction of the ordinary ray (o) is rotated from 0 deg (parallel to the x axis) to -45 deg. On the other hand, The polarization direction of the extraordinary ray (e) is rotated from +90 deg (parallel to the y axis) to

+45 deg. After this rotation, SP splits o and e horizontally and lines up the interferograms in the horizontal direction of the detector. Note that the interferograms can be aligned in the horizontal direction without HWP if the optical axis of the BPBS is in the +45 deg direction. However, prioritizing the ease of BPBS handling, we set the optical axis of the BPBS in the y axis direction.

In order to prevent misalignment between the four interferograms and the detector pixel array, each optical element must have an alignment accuracy of a few microns or better. This accuracy is achieved with the custom parts with leverage mechanisms and the commercial precision feed screws. The custom parts were manufactured at 3D Model Advanced Processing Center of Osaka Electro-Communication University.

Since this optical system contains many elements, the throughput is only about 70%. However, we can apply the original optical system with a small number of elements proposed in Tsukui+20 if multiple detectors are synchronized at 6.5 kHz in the future. In this case, the throughput is expected to be $\sim 90\%$, assuming a loss of 2% per surface.

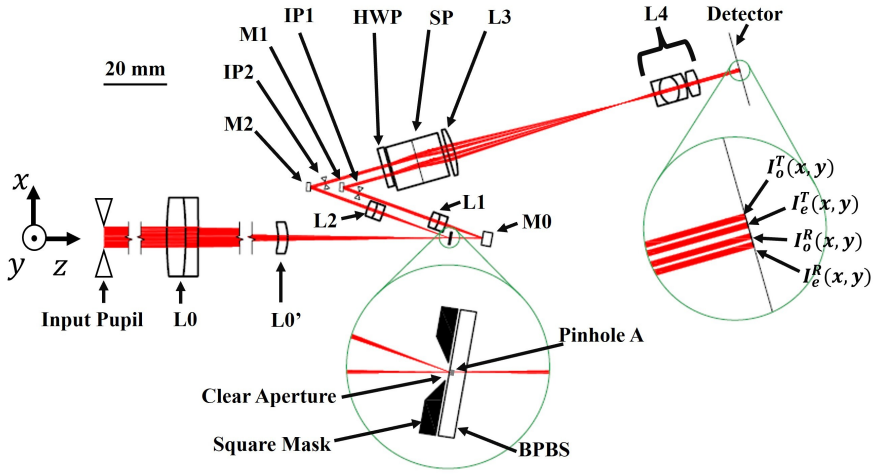


Figure 7.5: Layout of the b-PDI optical system. The x and y axes are perpendicular to the local axis of the beam and are horizontal and vertical, respectively. The polarization directions of the ordinary ray o and the extraordinary ray e are parallel to the x and y axes, respectively. The optical system includes a focusing lens L0, a correction lens L0', a square mask, the BPBS, a flat mirror M0, collimator lenses L1 and L2, flat mirrors M1 and M2, a half-wavelength plate HWP, a Savart plate SP, relay lenses L3 and L4, and a detector (Hamamatsu ORCA-Flash4.0V2). All are arranged in the horizontal plane. The BPBS is placed on the focal plane of L0. The center of the BPBS Pinhole A is on the axis of L0. The normal of the BPBS is inclined by 10 degrees to the axis of L0 in the $x-z$ plane. The inclination axis of the BPBS is parallel to the optical axis of TiO_2, which is parallel to the y-axis. The system from L0 to L4 fits on about the size of an A4 sheet of paper. Reprinted with permission from [58] ©Optica Publishing Group.

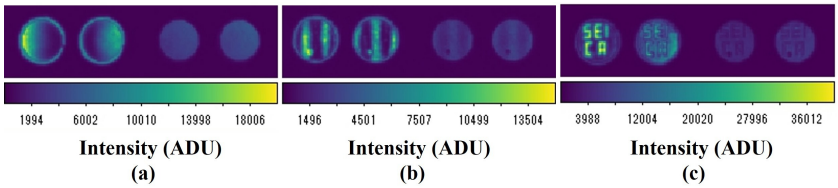


Figure 7.6: Examples of the interferograms captured with the manufactured b-PDI. (a) A tilted flat wavefront with P-V = 1.6 rad (200 nm). (b) A sinusoidal wavefront with 2 c/p and P-V = 1.2 rad (150 nm). (c) The figure "SEICA" generated with a deformable mirror whose stroke is $\sim$ 2 rad (250 nm). The interferograms in each image are $I_o^T(x,y), I_e^T(x,y), I_o^R(x,y)$, and $I_e^R(x,y)$ from left to right. Reprinted with permission from [58] ©Optica Publishing Group.

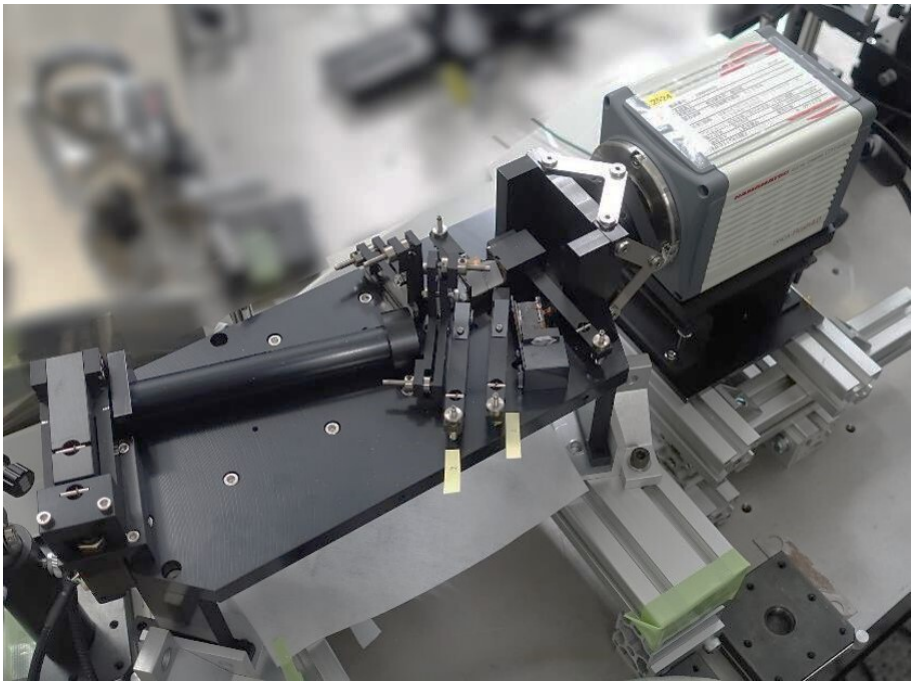


Figure 7.7: Photograph of the manufactured b-PDI optical system.

7.3 Summary of this chapter

This chapter has described the manufactured BPBS and b-PDI optical system. Among 192 manufactured BPBS samples, the one that produced interferograms with the best contrast was selected. The b-PDI with the BPBS sample (#03-40) is tested in detail in the next chapter.

Chapter 8

Laboratory Test and Discussion

This chapter describes the laboratory test of the b-PDI with the BPBS sample (#03-40). The test includes the evaluation of the systematic error σ_{sys} ($= \sqrt{\sigma_{rec}^2 + \sigma_{chrom}^2}$), the statistical error σ_{stat} ($= \sqrt{\sigma_{pn}^2 + \sigma_{rn}^2}$), and the dynamic range. This chapter then discusses the b-PDI's properties including the efficiency based on the evaluated σ_{stat}.

This chapter is based on Sections 4 and 5 of Tsukui et al. [58], to be published in Optics Continuum.

8.1 Laboratory test

8.1.1 Setup

We built a test bench (Fig. 8.1) to evaluate the b-PDI. In the test bench, the beam from the fiber source is collimated by CL and passes through the iris (6.5-mm diameter). RL1 and RL2 project the image of the iris onto the deformable mirror (DM; BMC 492-DM) at a magnification ratio of 1:1. The DM introduces phase aberrations to the beam. Then, the retractable FM1 switches the beam direction. When FM1 is moved out of the optical path, the beam directly enters the b-PDI. When FM1 is inserted into the optical path, the beam is reflected and enters the Shack-Hartmann WFS (SHWFS). In the SHWFS, RL3 and RL4 relay the beam to the WFS150-5C sensor head (Thorlabs), which contains a microlens array with 29×29 effective elements. The SHWFS is calibrated with FM2 ($\lambda/20$), which replace the DM. The microlens array, the detector of the b-PDI, and the DM are in the

conjugate plane.

We use two different fiber sources. For the SHWFS, a bright diode laser (Thorlabs CPS808S) with a wavelength of 800 nm is applied to reduce noise. The measurement error of the SHWFS with this source is 5×10^{-2} rad (7 nm) RMS, which is determined by measuring a spherical wavefront with a known radius of curvature. This error is primarily systematic. For the b-PDI, on the other hand, a tungsten-halogen lamp (Thorlabs SLS201L/M) is applied to simulate the weak stellar flux with a low coherency. The band-pass filters in the lamp unit limit the wavelengths to 800±100 nm.

When we inspect the point spread functions (PSFs) at the focal plane, we replace the SHWFS with a focal plane imager. The imager consists of an F/31 imaging lens and an sCMOS camera (Hamamatsu ORCA-Flash4.0V3). The configuration yields a PSF sampling of FWHM = 3.8 pix.

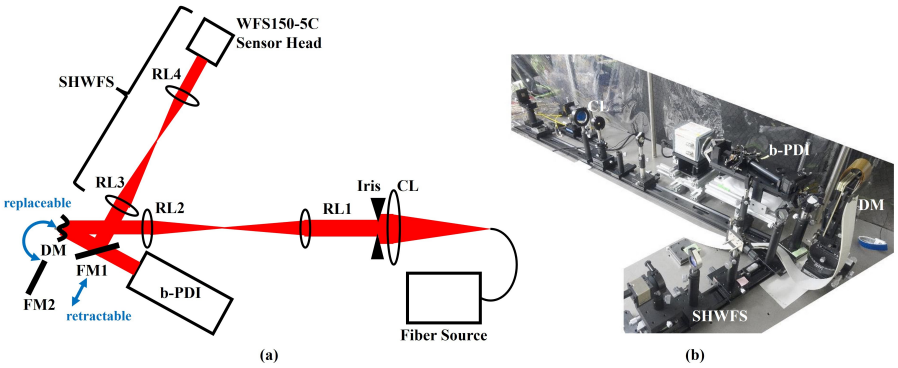


Figure 8.1: (a) Layout of the test bench. The size is 1.5 × 0.8 m. The test bench consists of a fiber source, a collimator lens CL, an iris (6.5-mm diameter), relay lenses RL1 and RL2, a deformable mirror (DM; BMC 492-DM), the b-PDI, a retractable flat mirror FM1 ($\lambda/20$), a flat mirror FM2 ($\lambda/20$), and a Shack-Hartmann wavefront sensor (SHWFS). The SHWFS consists of relay lenses RL3 and RL4, and a Thorlabs WFS150-5C sensor head with 29×29 subapertures. The DM has 492 (24×24) actuators. The b-PDI has 26×26 subapertures. (b) Photograph of the test bench viewed from the SHWFS side. Reprinted with permission from [58] ©Optica Publishing Group.

8.1.2 Measurement of the transmission/reflection coefficients

The phase reconstruction algorithms (Eqs. 5.21-5.23) requires the parameters expressed as the ratios of the transmission and reflection coefficients.

These parameters must be measured before the calculation. Specifically, the parameters in the Table 8.1 are to be measured.

In the measurement, a collimated beam with wavelengths of 800±100 nm was injected into the b-PDI. The BPBS was displaced laterally and placed so that Pinhole B or the bare TiO_2 crystal surface overlapped the Airy disk. The resulting pupil images were captured with the detector and photometrically measured. The parameters were calcurated by the ratio of the photometric values, as shown in the third row of Table 8.1. The IB and IC are the photometric values of the pupil images produced by Pinhole B and the bare TiO_2 crystal surface, respectively. The subscripts o and e denote the polarization direction, and the superscripts T and R denote the transmission and reflection path, respectively. The fourth row of Table 8.1 summarizes the results.

Table 8.1: Required parameters.

Parameter	t_o	t_e	r_o	r_e	s_o	s_e
Definition	$T^o_{\text{in}}/T^o_{\text{out}}$	$T^e_{\text{in}}/T^e_{\text{out}}$	$R^o_{\text{in}}/R^o_{\text{out}}$	$R^e_{\text{in}}/R^e_{\text{out}}$	$(R^o_{\text{out}}/T^o_{\text{out}})^2$	$(R^e_{\text{out}}/T^e_{\text{out}})^2$
Calculation	$\sqrt{IB^T_o/IC^T_o}$	$\sqrt{IB^T_e/IC^T_e}$	$\sqrt{IB^R_o/IC^R_o}$	$\sqrt{IB^R_e/IC^R_e}$	IC^R_o/IC^T_o	IC^R_e/IC^T_e
Measured value	1.01	1.03	0.89	0.84	0.22	0.30

8.1.3 Systematic error

We evaluated the systematic error σ_{sys} of the b-PDI by measuring known wavefronts. The systematic error was calculated from the difference between the known phase $\delta_0(x, y)$ and the phase $\delta_{rec,av}(x, y)$ measured by the b-PDI under a high S/N condition.

The wavefront under test was a sinusoidal wavefront with a spatial frequency of 8 c/p. The sinusoidal wavefront was generated with the DM and measured with the SHWFS to determine the phase. The tip/tilt components were removed from each of the 10 measured frames. The 10 frames were then converted to 26×26 pixels and averaged to be $\delta_0(x, y)$. The measured amplitude was $\simeq 0.24$ rad ($\simeq 30$ nm), which means P-V $\simeq 0.48$ rad ($\simeq 60$ nm). As mentioned earlier, the SHWFS has the systematic error of about 5×10^{-2} rad (7 nm) RMS.

We measured the wavefront with the b-PDI under the following conditions. By adjusting the integration time of the detector, the interferograms

with a high S/N ratio were captured. The average number of photons incident on one subaperture was 2.3×10^4 in the number of photo-electrons; the details of the photometry are described below. Ten frames of interferograms were converted to the phase $\delta_{rec}(x, y)$ with Eqs. 5.29-5.31 (the normal algorithm). The tip/tilt components were removed from each frame. The 10 frames of the phase were then averaged to be $\delta_{rec,av}(x, y)$. We evaluated the RMS value of the difference between $\delta_{rec,av}(x, y)$ and $\delta_0(x, y)$. Here, 21 subapertures were affected by a dead actuator on the DM and thus excluded.

Figure 8.2 shows the measured results and the difference. For the sinusoidal wavefront (8 c/p, P-V $\simeq 0.47$ rad), the RMS value of the difference was 8.4×10^{-2} rad (11 nm). This RMS value is the root-sum-square of the systematic errors of the SHWFS and the b-PDI. Considering that the SHWFS has a systematic error of $\sim 5 \times 10^{-2}$ rad (7 nm) RMS, the systematic error of the b-PDI is $\sim 7 \times 10^{-2}$ rad (9 nm) RMS. The systematic error could be reduced by a calibration with a flat wavefront.

The average photon number was estimated by the following photometry. By shifting the PSF core onto the bare TiO_2 crystal surface ~ 150 μm away from Pinhole A, we obtained the pupil images at the reflection path with less interference effect. The pupil images were captured at the same integration time as the above interferograms. The photometry of these images yielded the average number of photons (in the number of photo-electrons) in the reflection path. The average number of photons in the transmission path was estimated using the parameters s_o and s_e (Table 8.1). The sum of these numbers is the average number of photons in all the paths. Dividing this sum by the area ratio of the subaperture to the incident pupil, we got the above value (2.3×10^4 electrons). This value equals the average number of incident photons per subaperture when the throughput of the b-PDI optics (including the quantum efficiency of the detector) is 100%.

8.1.4 Statistical error

We evaluated the statistical error of the b-PDI by measuring sinusoidal wavefronts. Noise propagation theory predicts that the statistical error σ_{stat} [rad] of the b-PDI is expressed by the following equation:

$$\sigma_{stat}^2 = \sigma_{pn}^2 + \sigma_{rn}^2 \tag{8.1}$$

$$= \frac{C_p^2}{N_e} + \frac{C_r^2 N_r^2}{N_e^2}. \tag{8.2}$$

In the case of the zonal reconstruction in the b-PDI, N_e is the average number of photons (in the number of photo-electrons) incident on one sub-

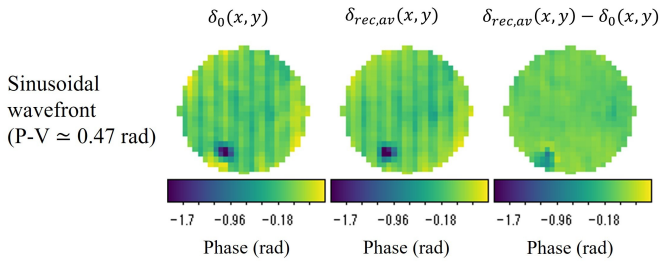


Figure 8.2: Comparison between $\delta_{rec,av}(x,y)$ and $\delta_0(x,y)$. The map $\delta_{rec,av}(x,y)$ is the average of the 10 frames of the phase, which are measured with the b-PDI under a high S/N condition. The map $\delta_0(x,y)$ is the average of the 10 frames of the phase, which are measured with the SHWFS under a high S/N condition. Reprinted with permission from [58] ©Optica Publishing Group.

aperture. The terms C_p and C_r are the coefficients of error propagation from photon noise and readout noise, respectively. The term N_r is the readout noise; $N_r = 1.6$ electrons RMS in this experiment [36].

As the prediction, we ran simulations using the numerical model (Section 6.1.1), which reflected the actual Pinhole A (#03-40) geometry shown in Table 7.1. The values of C_p and C_r were predicted from the simulations with photon noise or readout noise at $N_e > 300$. Table 8.2 shows the result with sinusoidal wavefronts. The spatial frequencies of the wavefronts were 2, 4, and 8 c/p. The amplitudes of the wavefronts were 0.16 rad (20 nm), which means P-V = 0.32 rad (40 nm).

Table 8.2: Predicted C_p and C_r from the simulations.

Spatial frequency of wavefront	C_p	C_r
2 c/p	1.0	3.1
4 c/p	1.0	3.1
8 c/p	1.0	3.0

In the experiment, the sinusoidal wavefronts are shaped with the DM. We measured the wavefronts with the b-PDI under the following conditions. By adjusting the integration time of the detector, N_e was set in a range of $24 - 2.4 \times 10^4$. The N_e was estimated in the same way as in Section 8.1.3 and scaled by the integration time. At each integration time, 100

frames of interferograms were captured. However, 10 frames were captured at $N_e = 2.4 \times 10^4$ electrons. The phase $\delta_{rec}(x, y)$ was calculated at each subaperture (x, y) with Eqs. 5.29-5.31 (the normal algorithm). This calculation used two types of numerical masks: fixed and low-count masks. The fixed masks excluded the eight subapertures that hit the beam's edges and the 12 subapertures that overlap the dead actuator on the DM. The low-count masks excluded subapertures with

$$\frac{I_o^R(x, y) + I_e^R(x, y)}{2} < 2 \text{ ADU } (\simeq 1 \text{ electron}). \tag{8.3}$$

This operation excludes physically meaningless count values. The tip/tilt components were subtracted from the calculated phase $\delta_{rec}(x, y)$.

The experimental values of the statistical error σ_{exp} were calculated from $\delta_{rec}(x, y)$. First, $\delta_{rec}(x, y)$ at $N_e = 2.4 \times 10^4$ electrons was averaged over 10 frames and defined as the reference phase $\delta_{Ref}(x, y)$. Next, the difference between $\delta_{Ref}(x, y)$ and $\delta_{rec}(x, y)$ at $N_e < 2.4 \times 10^4$ electrons was taken, and its RMS value was calculated. This RMS value is σ_{exp}. Note that σ_{exp} contains only the statistical errors because the difference canceled the systematic errors. The median, maximum and minimum values of σ_{exp} were selected from the 100 frames at each integration time.

We compared the experimental results σ_{exp} with its predicted value σ_{stat} (Eq. 8.2). Figure 8.3 shows the results. The experimental σ_{exp} and its predicted value σ_{stat} are consistent in the region $N_e \gtrsim 100$. At $N_e = 60$, the experimental σ_{exp} exceeded the predicted σ_{stat}. This implies nonlinear error propagation, with effective C_p and C_r being larger than the values in Table 8.2, respectively. These elevated values were also observed in the simulations. For example, for the 4 c/p sinusoidal wavefront, the simulated effective C_p and C_r are 1.4 and 4.1 at $N_e = 60$, respectively. These simulated values are consistent with the experimental value σ_{exp} at $N_e = 60$. On the other hand, the experimental value σ_{exp} was lower than the predicted value σ_{stat} at $N_e = 24$. This is because the low-count masks eliminated the subapertures with large noise and small count values below 1 electron, suppressing the overall error (see Fig. 8.3 (d)). In this region, the effective C_p and C_r are smaller than the values in Table8.2, respectively. The low-count masks removed $\sim 30\%$ of the subapertures at $N_e = 24$ and $\sim 5\%$ at $N_e = 60$.

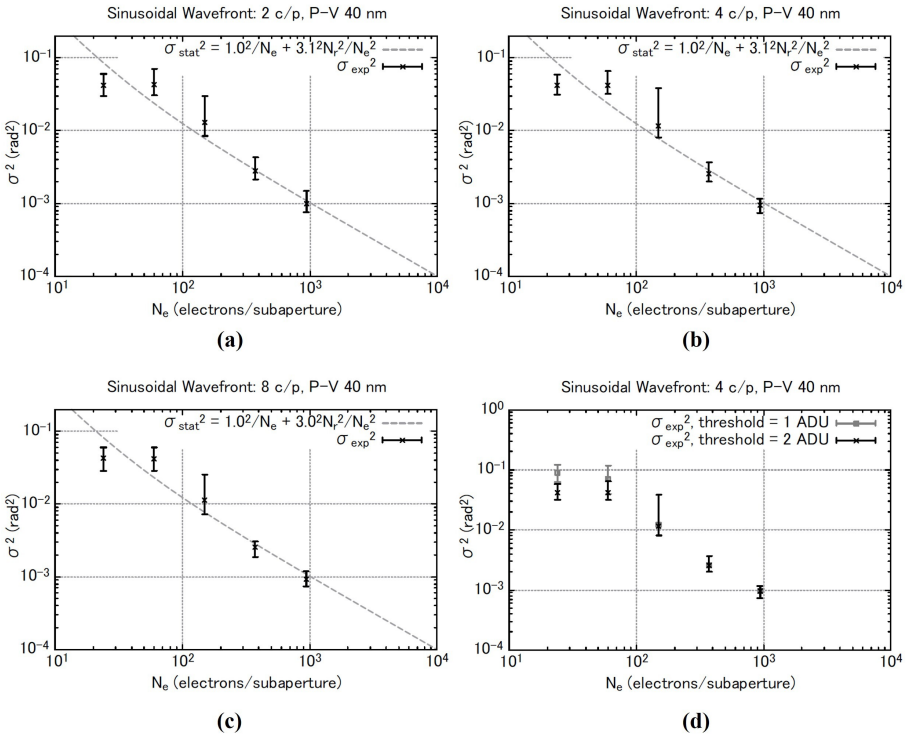


Figure 8.3: (a)(b)(c) Comparison of the experimental value σ_{exp} of the statistical error with its predicted value σ_{stat}. The horizontal axis is the average number of photons per subaperture N_e (in the number of photo-electrons). The graphs show results with the sinusoidal wavefronts (2, 4, and 8 c/p), as noted in each panel. The gray dashed line represents the predicted σ_{stat}, the black dots are the median of σ_{exp}, and the error bars are their maximum and minimum values. The N_r represents the readout noise; $N_r = 1.6$ electrons RMS in this experiment. (d) Experimental values σ_{exp} of the statistical error with different thresholds for the low-count masks. For the black dots, the low-count masks are applied at < 2 ADU ($\simeq 1$ electron). For the gray dots, the low-count masks are applied at < 1 ADU ($\simeq 0.5$ electron). Reprinted with permission from [58] ©Optica Publishing Group.

8.1.5 Limit of dynamic range

We evaluated the dynamic range of the b-PDI, using the sinusoidal wavefronts generated with the DM. The spacial frequency was 1, 2, and 4 c/p.

First, we measured the wavefronts with the SHWFS. Ten frames were

measured. The tip/tilt components were removed from each frame. The 10 frames were then converted to 26×26 pixels. The measured phase of the SHWFS at frame number j was defined as $\delta_{SH,j}(x, y)$.

We then measured the wavefronts with the b-PDI. By adjusting the integration time, the number of photons was set to $N_e \gtrsim 1.2 \times 10^4$ electrons. Ten frames of interferograms were captured. The phase was calculated at each subaperture with Eqs. 5.29-5.31 (the normal algorithm). The measured phase of the b-PDI at frame number j was defined as $\delta_{rec,j}(x, y)$.

We then calculated the magnification factor M_j of the b-PDI relative to the SHWFS at frame number j. The optimal value as M_j minimizes the following evaluation function $D(m)$:

$$D(m) = \Sigma_{x,y} \left\{ \delta_{rec,j}(x, y) - m\delta_{SH,j}(x, y) \right\}^2 . \tag{8.4}$$

In other words, M_j satisfies the following equation:

$$\left. \frac{\partial D(m)}{\partial m} \right|_{m=M_j} = 0. \tag{8.5}$$

Therefore, M_j is defined as

$$M_j = \frac{\Sigma_{x,y}\, \delta_{rec,j}(x, y)\, \delta_{SH,j}(x, y)}{\Sigma_{x,y}\, \delta_{SH,j}(x, y)^2}. \tag{8.6}$$

The 16 or 25 subapertures were affected by the dead actuator of the DM and thus numerically masked. The median, maximum and minimum values of M_j were selected from the 10 frames.

Finally, we captured the PSF with the focal plane imager to calculate the Strehl ratio. The calculation was based on Method Six described by Roberts et al. [60].

Figures 8.4 and 8.5 show the result with the normal and iterative algorithms, respectively. Since the arctangent function atan2 does not cause ambiguity in the range of $1\lambda_C$ in P-V, $M_j > 0$ was expected in the same range. However, we observed $M_j = 0$ at P-V $< 1\lambda_C$, and the measured phase was almost flat. The $M_j = 0$ occurred when the Strehl ratio was ~ 0 at any spatial frequency. This is because the reference beam disappears, and the contrast of the interferogram takes ~ 0 at the Strehl ratio ~ 0. At this point, interferometry can not be performed. This effect sets a practical upper limit of the dynamic range. Thus, the dynamic range is P-V $< 0.6\lambda_C - 1.0\lambda_C$ for the sinusoidal wavefronts with $4 - 1$ c/p. In the dynamic range, the wavefront phase is calculated without being affected by the ambiguity of the equation.

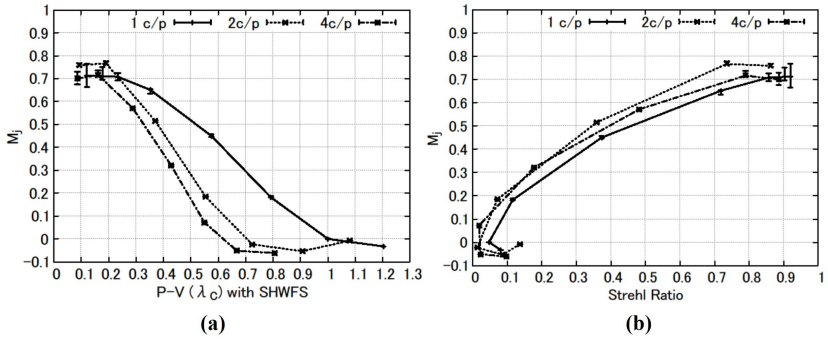


Figure 8.4: (a) Magnification factor M_j of the b-PDI with the normal algorithm, plotted against the P-V value of the phase measured with the SHWFS. (b) Magnification factor M_j plotted against the Strehl ratio of the PSF. In each panel, the black dots represent the median of M_j, and the error bars are their maximum and minimum values. The line types identify the spatial frequency of the wavefronts: solid line for 1 c/p, dashed line for 2 c/p, and dashed-and-dotted line for 4 c/p. Reprinted with permission from [58] ©Optica Publishing Group.

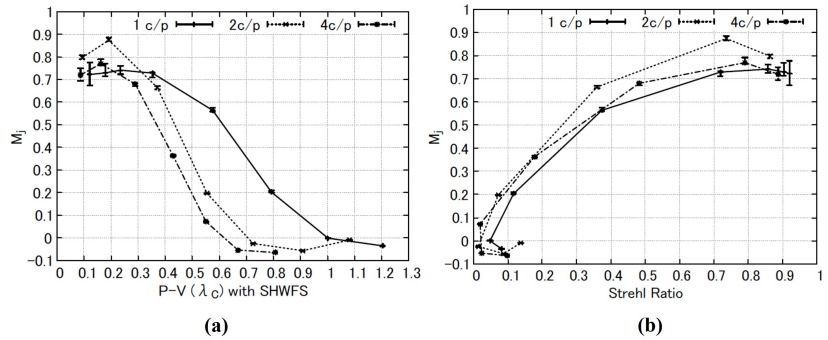


Figure 8.5: (a) Magnification factor M_j of the b-PDI with the iterative algorithm (five iterations), plotted against the P-V value of the phase measured with the SHWFS. (b) Magnification factor M_j plotted against the Strehl ratio of the PSF.

8.1.6 Linearity in practical cases

For a more practical evaluation, the dynamic range characterized by linearity is worth investigating. Natural atmospheric wavefronts in practical use are less prone to degrade the Strehl ratio than sinusoidal wavefronts because natural ones include multiple spatial frequencies. Thus, non-sinusoidal wave-

fronts are more suitable for evaluating the b-PDI behavior with large aberrations. We evaluated the linearity within P-V $< 1\lambda_C$, using a non-sinusoidal wavefront shaped by the DM. The wavefront under test was expressed by the Zernike polynomial Z_{20} (pentafoil), which does not yield $SR = 0$ within the range (P-V $< 1\lambda_C$).

First, we measured the wavefront with the SHWFS in the same way as the previous section to obtain $\delta_{SH,j}(x, y)$ (26×26 pixels). The RMS values of $\delta_{SH,j}(x, y)$ were calculated, and the median value was selected to be RMS_{input}. This was regarded as the RMS value of the input wavefront phase. Note that $RMS_{input} \simeq 0.7$ rad corresponds to P-V $= 1\lambda_C$.

We then measured the wavefront with the b-PDI. We used two fiber sources: the tungsten-halogen lamp (800 ± 100 nm) and the diode laser (800 nm). The latter was used to evaluate the effect of coherency and to compare the b-PDI linearity with the FPyWFS evaluated with a laser source [61]. The number of photons was set to $N_e \gtrsim 1.2 \times 10^4$ electrons. With each fiber source, ten frames of interferograms were captured, and $\delta_{rec,j}(x, y)$ were calculated with both the normal and iterative algorithms. The iterative algorithm required five iterations to converge the calculated phase. The RMS values of $\delta_{rec,j}(x, y)$ were calculated, and the median values were selected to be RMS_{output}.

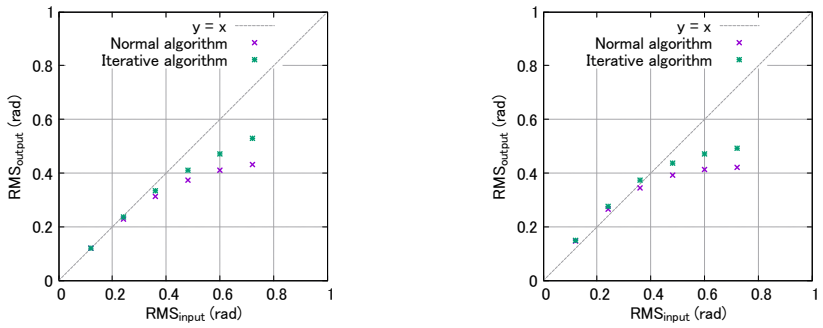
Finally, we compared RMS_{output} with RMS_{input}. Figure 8.6 shows the result.

The normal algorithm lets the b-PDI behave linearly within $RMS_{input} \lesssim$ 0.4 rad. The algorithm underestimated the input phase aberrations with $RMS_{input} \gtrsim 0.4$ rad. This is partially due to the mismatch of the assumed f in the calculation and the attenuated reference beam. The iterative algorithm improved the linearity by compensating for the mismatch.

The diode laser (800 nm) yielded better linearity compared to the tungsten-halogen lamp (800 ± 100 nm). This reflects the improved contrasts of the interferograms due to better coherency. With the iterative algorithm, the b-PDI is close to linear within $RMS_{input} \lesssim 0.5$ rad. In contrast, the simulation shown in Appendix A predicted linear response within $RMS_{input} \lesssim 0.7$ rad (P-V $\leq 1\lambda_C$) with a single wavelength (800 nm). The difference between the experimental and simulated results is due to factors not considered in the simulation, such as scattered light, stray light, and vibration.

The experimental result with the diode laser can be compared with the FPyWFS performance evaluated with a 635-nm laser source [61], shown in Fig. 8.7. The response curve of the FPyWFS with an MVM-like algorithm deviates from linear at $RMS_{input} \simeq 0.3$ rad (P-V$\simeq \lambda_C/2$). This reflects the limit of the approximation: $\sin x = x$ (Eq. 4.14). With $RMS_{input} \gtrsim 0.3$

rad, RMS_{output} is underestimated and saturates. Thus, the dynamic range with a linear response is $RMS_{input} \lesssim 0.3$ rad. On the other hand, the b-PDI's response with the iterative algorithm (Fig. 8.6(b)) is close to linear in $RMS_{input} \lesssim 0.5$ rad.



(a) **(b)**

Figure 8.6: The response of RMS_{output} against RMS_{input} with (a) the tungsten-harogen lamp (800 ± 100 nm) and (b) the diode laser (800 nm).

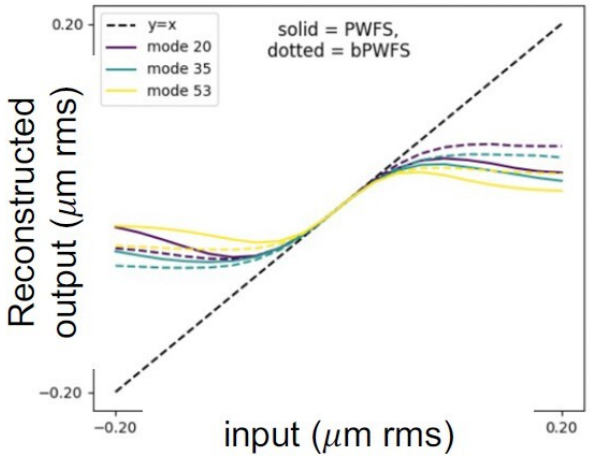


Figure 8.7: The response curves of FPyWFS, evaluated with a 635-nm laser source and an MVM-like algorithm. The figure is cited from Gerard et al. (2022) [61]. The purple solid line represents the response curve for Z_{20} wavefront. 0.20 μm rms corresponds to 2.0 rad RMS.

8.2 Discussion

8.2.1 Statistical error and efficiency

The coefficient C_p describes the contribution of photon noise to the statistical error. The experiments showed that the manufactured b-PDI has C_p of $\gtrsim 1.0$, which is $\gtrsim 1.4$ times larger than that of a Zernike WFS [48]. This means that the b-PDI is $\gtrsim 1.4$ times more sensitive to photon noise than a Zernike WFS. Thus, the sensitivity factor β_p of the manufactured b-PDI is $\gtrsim 1.4$. In other words, the efficiency of the b-PDI is comparable to that of a fixed pyramid WFS.

The further enhancement of the b-PDI efficiency is one of the future works, considering the simulated efficiency in Section 6.1.2. For example, the following are to be investigated. (1) Further optimization of the pinhole depth d under chromatic effects. (2) Modifying the reconstruction algorithm to avoid the strong error propagation at $N_e \sim 60$. (3) Investigating other limiting factors such as scattered light.

8.2.2 High-speed capabilities

The high-speed capabilities include a low calculation cost and a small readout region.

The calculation cost of the b-PDI is relatively low. The cost is $\mathcal{O}(N_{SA})$, where N_{SA} is the total number of the subapertures. This is as low as the cost of a Zernike WFS, and that of the fastest algorithm of a fixed pyramid WFS [43].

The b-PDI has a relatively small readout region in the vertical direction on a detector. The size in the direction is N_{1d} pixels, where N_{1d} is the number of subapertures filling the aperture diameter. A Zernike WFS and a fixed pyramid WFS require N_{1d} and $> 2N_{1d}$ pixels in the direction, respectively. The fewer pixels in the vertical direction of sCMOS detectors reduce the time for readout and yield a higher frame rate.

8.2.3 Dynamic range

The b-PDI has a relatively large dynamic range in two ways of characterizations: (1) the range in which the reconstruction equation is applicable and (2) the range in which the WFS responds linearly. As for (1), the b-PDI's reconstruction equations (Eqs. 5.29-5.31) calculates the tangent (sine/cosine) of the phase aberration δ. Thus the equation does not cause ambiguity in the range $|\delta| < \pi$, which corresponds to P-V $< 1\lambda_C$. This is

about two times larger than the range of other phase sensors (ZWFS and FPyWFS). We do not willingly use phase unwrapping algorithms to reduce the calculation cost. However, if the cost is not an issue, there is room to use the algorithms to overcome the limitation of the atan2 function. As for (2), the b-PDI responded linearly in $RMS_{input} \lesssim 0.5$ rad with the iterative reconstruction algorithm and the coherent beam. This is larger than the range of FPyWFS evaluated by Gerard et al. [61] with a coherent beam.

Note that the b-PDI also has upper limit of the dynamic range set by the Strehl ratio, as described in Section 8.1.5. This is because the reference beam disappears at a Strehl ratio of ~ 0, resulting in unsuccessful interferometry. This effect is inevitable and limits the dynamic range of any PDI.

8.2.4 Low-count masks

The numerical low-count masks reduce the statistical errors when the b-PDI captures a low number of incident photons ($N_e \sim 24$ electrons). In the experiment, the subapertures with signals equivalent to $\lesssim 1$ electron were masked. The masked subapertures were $\sim 30\%$ of the total at $N_e = 24$ electrons. In the future, detectors with smaller readout noise (< 1 electron) will reduce the number of the masked subapertures. Simulations predict that the fraction of the subapertures masked when the readout noise is 0.3 and 0 electrons is 21% and 12%, respectively. In the practical use in a closed-loop ExAO, the unmeasurable phase values in the masked subaperture can be compensated for by the phase values measured at the previous loop.

8.2.5 Scalability and limiting factor

ExAO in a larger telescope requires more subapertures and higher resolution of the interferograms because D/r_0 increases. The scalability of the b-PDI resolution is limited by the detector size; the diameter of each interferogram must be smaller than a quarter of the detector width.

8.3 Summary of this chapter

This chapter has described the result of the laboratory test and discussed the properties of the b-PDI. Table 8.3 compares the manufactured b-PDI's performance with the simulated and target performance (Table 5.1). Table 8.4 compares the manufactured b-PDI with other phase sensors (Table 4.1).

Table 8.3: Comparison of the manufactured b-PDI's performace with the target and simulated performance. β_p: sensitivity factor [27]. N_{1d}: 1-dimensional number of the effective subapertures. λ_C: center of the sensing wavelength band. $\Delta\lambda$: width of the sensing wavelength band.

	Target	Simulated	Manufactured
Type	Phase sensor	Phase sensor	Phase sensor
β_p	≤ 1.4	$\simeq 1.1$	$\gtrsim 1.4$
Sampling rate	6.5 kHz [i]	6.5 kHz	7.9 kHz [iii]
Calculation cost	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA})$
Number of detector pixels (vertical)	N_{1d}	N_{1d}	N_{1d}
N_{1d}	≥ 24 [i]	24	26
Dynamic range	$1\lambda_C$	$1\lambda_C$ [ii]	$1\lambda_C$ [ii]
λ_C	800 nm [i]	800 nm	800 nm

[i] Assuming the SEICA instrument [17].
[ii] The range in which the reconstruction equation is applicable without ambiguity.
[iii] Maximum frame rate of ORCA-Flash 4.0V2 CMOS camera (free running mode) with 26 vertical pixels [36].

Table 8.4: Comparison of WFSs. C_p: coefficient of error propagation from photon noise. β_p: sensitivity factor [27]. λ_C: center of the sensing wavelength band. N_{SA}: number of all the effective subapertures. N_{1d}: 1-dimensional number of the effective subapertures.

	ZWFS	FPyWFS	b-PDI (simulated)	b-PDI (Manufactured)
Type	Phase	Phase	Phase	phase
C_p	$\sqrt{0.5}$ [i]	N/A	0.77	$\gtrsim 1.0$
β_p	1.0	$\sqrt{2}$	$\simeq 1.1$	$\gtrsim 1.4$
Calculation cost	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA}) - \mathcal{O}(N_{SA}^2)$	$\mathcal{O}(N_{SA})$	$\mathcal{O}(N_{SA})$
Number of detector pixels (vertical)	N_{1d}	$> 2N_{1d}$	N_{1d}	N_{1d}
Dynamic range	$< 0.5\lambda_C$ [ii]	$\lesssim 0.5\lambda_C$ [ii]/ $\lesssim 0.3$ rad RMS [iii]	$1\lambda_C$ [ii]/ $\lesssim 0.7$ rad RMS [iii]	$1\lambda_C$ [ii]/ $\lesssim 0.5$ rad RMS [iii]

[i] From N'Diaye et al. 2013 [48].
[ii] The range in which the reconstruction equation is applicable without ambiguity.
[iii] The range in which the WFS responds linearly with the Z_{20} wavefront and coherent beams.

Chapter 9

Conclusion

Direct observation of exoplanets is crucial for solving the following two problems: clarifying the mechanisms of planet formation and searching for extraterrestrial biological activities. Potential targets have small angular separations ($\sim 0.01 - 0.1$ arcsec) from their host stars and small planet-to-star contrasts ($\sim 10^{-7}$). Thus, direct observation requires a high angular resolution and high contrast. Such observation with ground-based telescopes is affected by the Earth's atmospheric turbulence. This is because the turbulence causes wavefront aberrations, which scatter the host star's light to degrade the angular resolution and contrast.

Therefore, a highly accurate wavefront correction is required to obtain a 10^{-7}-level final contrast. Such correction can be achieved with extreme adaptive optics (ExAO), which has ~ 40000 measurement/correction points in a 30-m telescope aperture and runs at ~ 5 kHz. ExAO requires a high-performance WFS with the following properties:

- high efficiency: a small measurement error with a limited number of photons,
- high-speed capabilities: low calculation cost and a small readout region for wavefront sensing,
- a large dynamic range.

As for high efficiency, ExAO favors phase sensors, such as a fixed pyramid WFS and a Zernike WFS. However, These current phase sensors have room to improve high-speed capabilities and dynamic ranges.

This thesis has described the development of a new phase sensor named b-PDI. The b-PDI utilizes birefringent crystal (TiO_2 in this thesis) as its key

optical element: the BPBS. The presented simulations determined the optimal design of the BPBS. The BPBS and optical systems were manufactured and tested in the laboratory. The following properties were achieved.

- high efficiency (the sensitivity factor $\beta_p \gtrsim 1.4$) comparable to other phase sensors,
- high-speed capabilities with lowest level of calculation cost and smallest level of readout region (in the vertical direction) for wavefront sensing,
- a larger dynamic range in two ways of characterizations: (1) the range in which the reconstruction equation is applicable without ambiguity and (2) the range in which the WFS responds linearly.

Table 8.4 shows the detailed comparison.

Chapter 10

Ongoing and Future Works

10.1 ExAO experiment

As a part of the development of SEICA instrument, the laboratory demonstration of ExAO based on the b-PDI is currently in progress. The experiment utilizes the ExAO (only Tweeter) testbed consisting of:

- WFS: b-PDI,
- DM: BMC 492-DM,
- RTC: Xilinx Zynq UltraScale+ MPSoC ZCU102 (Field programmable gate array; FPGA).

The following operation has been established so far.

- The RTC receives the signal from the b-PDI and calculates the wavefront aberration.
- An arbitrary actuator of the DM moves by the commands from the RTC.

Further debug and optical alignment are required to establish the closed-loop ExAO operation. The FPGA-based high-speed RTC is expected to yield $F_{AO} \sim 7$ kHz.

10.2 Contribution to TMT

The experimental results of the b-PDI was introduced in "TMT の次期装置実現に向けた開発ロードマップ" (The development roadmap for realization of

TMT's next instruments) [62] as an achievement in the development of the high contrast instrument of TMT. Further contribution to TMT is expected.

Acknowledgement

I would like to express my appreciation and thanks to those who supported and contributed to this study. Associate professor Mikio Kurita, who is my supervisor, encouraged me and gave me many important comments on this study and my career. Assistant professor Masaru Kino always helped me with technical issues and led the design of the manufactured b-PDI optical system. Assistant professor Kodai Yamamoto gave me a lot of chance of discussion and contributed to the setup of the test bench. Dr. Akihiro Matsutani, Ms. Mina Sato, Ms. Mie Tohnishi, Ms. Miho Fujimoto and Ms. Yayoi Hoshino at Tokyo Institute of Technology cooperated in the BPBS manufacturing and gave me detailed instructions on microprocessing even in the difficult times of the COVID-19 pandemic. Former professor Tetsuya Nagata and the members of the instrumentation group gave me chance of discussion. Lecturer Naoshi Murakami at Hokkaido University gave me advice on the simulations. Professor Masatsugu Iribe at Osaka Electro-Communication University (OECU) gave me advice on control engineering and my career. Ms. Yukiko Nakanishi, Ms. Erika Touma, Ms. Noriko Yamamoto, Ms. Miyuki Kawamoto, Ms. Ayako Nagaoka, Ms. Eri Okamoto and Ms. Keiko Ohara at Kyoto University helped me with travel procedures and budget management. I would also like to express my appreciation and thanks to the members of the development team of Seimei 3.8-m segmented telescope, Kyoto Prefectural Technology Center, 3D Model Advanced Processing Center of OECU, MIZOJIRI OPTICAL CO., LTD., SHINKOSHA CO., LTD., TOSHIMA Manufacturing CO., LTD., JASCO Corporation, Musashino Denshi, INC., KYODO INTERNATIONAL, INC., OKAMOTO OPTICS, INC., Japan Cell CO., LTD., and Orange Gazai.

I would like to express my special thanks to my family members. Their support is indispensable for the execution of this study.

This study is also supported by Japan Science and Technology Agency (JPMJSP2110) and Japan Society for the Promotion of Science (JP16H03957, JP18H03719).

Bibliography

[1] Michel Mayor and Didier Queloz. A Jupiter-mass companion to a solar-type star. Nature, 378(6555):355–359, 1995.

[2] NASA. Exoplanet exploration. `https://exoplanets.nasa.gov`. Accessed on 2022-10-20.

[3] Gerard P Kuiper. On the origin of the solar system. Proceedings of the National Academy of Sciences, 37(1):1–14, 1951.

[4] AW Harris. The formation of the outer planets. In Lunar and Planetary Science Conference, volume 9, pages 459–461, 1978.

[5] Hiroshi Mizuno, Kiyoshi Nakazawa, and Chushiro Hayashi. Instability of a gaseous envelope surrounding a planetary core and formation of giant planets. Progress of Theoretical Physics, 60(3):699–710, 1978.

[6] Hiroshi Mizuno. Formation of the giant planets. Progress of theoretical physics, 64(2):544–557, 1980.

[7] David S Spiegel and Adam Burrows. Spectral and photometric diagnostics of giant planet formation scenarios. The Astrophysical Journal, 745(2):174, 2012.

[8] Olenka Hubickyj, Peter Bodenheimer, and Jack J Lissauer. Accretion of the gaseous envelope of Jupiter around a 5–10 Earth-mass core. Icarus, 179(2):415–431, 2005.

[9] Mark S Marley, Jonathan J Fortney, Olenka Hubickyj, Peter Bodenheimer, and Jack J Lissauer. On the luminosity of young Jupiters. The Astrophysical Journal, 655(1):541, 2007.

[10] Michaël Gillon, Amaury HMJ Triaud, Brice-Olivier Demory, Emmanuël Jehin, Eric Agol, Katherine M Deck, Susan M Lederer, Julien De Wit,

Artem Burdanov, James G Ingalls, et al. Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1. Nature, 542(7642):456–460, 2017.

[11] George Evelyn Hutchinson. The biochemistry of the terrestrial atmosphere. The earth as a planet, page 371, 1954.

[12] Dian R Hitchcock and James E Lovelock. Life detection by atmospheric analysis. Icarus, 7(1-3):149–159, 1967.

[13] Olivier Guyon. Extreme adaptive optics. Annual Review of Astronomy and Astrophysics, 56:315–355, 2018.

[14] Victoria S Meadows. Reflections on O2 as a biosignature in exoplanetary atmospheres. Astrobiology, 17(10):1022–1052, 2017.

[15] Victoria S Meadows, Christopher T Reinhard, Giada N Arney, Mary N Parenteau, Edward W Schwieterman, Shawn D Domagal-Goldman, Andrew P Lincowski, Karl R Stapelfeldt, Heike Rauer, Shiladitya DasSarma, et al. Exoplanet biosignatures: understanding oxygen as a biosignature in the context of its environment. Astrobiology, 18(6):630–662, 2018.

[16] Edward W Schwieterman, Victoria S Meadows, Shawn D Domagal-Goldman, Drake Deming, Giada N Arney, Rodrigo Luger, Chester E Harman, Amit Misra, and Rory Barnes. Identifying planetary biosignature impostors: spectral features of CO and O4 resulting from abiotic O2/O3 production. The Astrophysical Journal Letters, 819(1):L13, 2016.

[17] Taro Matsuo, Naoshi Murakami, Takayuki Kotani, Hajime Kawahara, Noriaki Natsume, Masaru Kino, Kodai Yamamoto, Hiroaki Imada, Mikio Kurita, Masatsugu Iribe, et al. High-contrast planet imager for Kyoto 4m segmented telescope. In Ground-based and Airborne Instrumentation for Astronomy V, volume 9147, pages 686–705. SPIE, 2014.

[18] Stefano Casertano, Mario G Lattanzi, Alessandro Sozzetti, Alessandro Spagna, Sylvie Jancart, Roberto Morbidelli, Renato Pannunzio, Dimitri Pourbaix, and Didier Queloz. Double-blind test program for astrometric planet detection with Gaia. Astronomy & Astrophysics, 482(2):699–729, 2008.

[19] Olivier Guyon, Benjamin Mazin, Michael Fitzgerald, Dimitri Mawet, Christian Marois, Andy Skemer, Julien Lozi, and Jared Males. Wavefront control architecture and expected performance for the TMT Planetary Systems Imager. In Adaptive Optics Systems VI, volume 10703, pages 327–337. SPIE, 2018.

[20] J Hardy. Adaptive Optics for Astronomical Telescopes, 1998.

[21] Andrey Nikolaevich Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. Cr Acad. Sci. URSS, 30:301–305, 1941.

[22] François Roddier. V the effects of atmospheric turbulence in optical astronomy. In Progress in optics, volume 19, pages 281–376. Elsevier, 1981.

[23] AM Obukhov. Temperature field structure in a turbulent flow. Izv. Acad. Nauk SSSR Ser. Geog. Geofiz, 13:58–69, 1949.

[24] David L Fried. Statistics of a geometric representation of wavefront distortion. JoSA, 55(11):1427–1435, 1965.

[25] Robert J Noll. Zernike polynomials and atmospheric turbulence. JOsA, 66(3):207–211, 1976.

[26] Lisa Poyneer, Marcos van Dam, and Jean-Pierre Véran. Experimental verification of the frozen flow atmospheric turbulence assumption with use of astronomical adaptive optics telemetry. JOSA A, 26(4):833–846, 2009.

[27] Olivier Guyon. Limits of adaptive optics for high-contrast imaging. The Astrophysical Journal, 629(1):592, 2005.

[28] Richard Hudgin. Wave-front compensation error due to finite corrector-element size. JOSA, 67(3):393–395, 1977.

[29] Bengt Edlén. The refractive index of air. Metrologia, 2(2):71, 1966.

[30] Olivier Guyon, EA Pluzhnik, MJ Kuchner, B Collins, and ST Ridgway. Theoretical limits on extrasolar terrestrial planet detection with coronagraphs. The Astrophysical Journal Supplement Series, 167(1):81, 2006.

[31] Christian Marois, David Lafreniere, René Doyon, Bruce Macintosh, and Daniel Nadeau. Angular differential imaging: a powerful high-contrast imaging technique. The Astrophysical Journal, 641(1):556, 2006.

[32] Jun Nishikawa. The coherent differential imaging on speckle area nulling (CDI-SAN) method for high-contrast imaging under speckle variation. The Astrophysical Journal, 930(2):163, 2022.

[33] Pascal J Bordé and Wesley A Traub. High-contrast imaging from space: speckle nulling in a low-aberration regime. The Astrophysical Journal, 638(1):488, 2006.

[34] Mikio Kurita, Masaru Kino, Fumihide Iwamuro, Kouji Ohta, Daisaku Nogami, Hideyuki Izumiura, Michitoshi Yoshida, Kazuya Matsubayashi, Daisuke Kuroda, Yoshikazu Nakatani, et al. The Seimei telescope project and technical developments. Publications of the Astronomical Society of Japan, 72(3):48, 2020.

[35] TMT-PSI team. Wavefront control architecture and expeted performance for the TMT Planetary Systems Imager. `https://www.naoj.org/staff/guyon/publications/2018/2018-06-11_SPIE/TMTpsi/TMTPSI-WFC.pdf`. Accessed on 2022-12-02.

[36] Hamamatsu Photonics K.K. ORCA-Flash4.0 V2 Digital CMOS Camera C11440-22CU/C11440-22CU01 Instruction manual Ver.1.7, 2016.

[37] Francois Roddier. Curvature sensing and compensation: a new concept in adaptive optics. Applied Optics, 27(7):1223–1225, 1988.

[38] Hideki Takami, Naruhisa Takato, Yutaka Hayano, Masanori Iye, Shin Oya, Yukiko Kamata, Tomio Kanzawa, Yosuke Minowa, Masashi Otsubo, Koji Nakashima, et al. Performance of Subaru Cassegrain adaptive optics system. Publications of the Astronomical Society of Japan, 56(1):225–234, 2004.

[39] Olivier Guyon. High sensitivity wavefront sensing with a nonlinear curvature wavefront sensor. Publications of the Astronomical Society of the Pacific, 122(887):49, 2009.

[40] Ben C Platt and Roland Shack. History and principles of Shack-Hartmann wavefront sensing, 2001.

[41] Roberto Ragazzoni. Pupil plane wavefront sensing with an oscillating prism. Journal of modern optics, 43(2):289–293, 1996.

[42] Christophe Vérinaud. On the nature of the measurements provided by a pyramid wave-front sensor. Optics Communications, 233(1-3):27–38, 2004.

[43] Iuliia Shatokhina, Victoria Hutterer, and Ronny Ramlau. Review on methods for wavefront reconstruction from pyramid wavefront sensor data. Journal of Astronomical Telescopes, Instruments, and Systems, 6(1):010901, 2020.

[44] Jianxin Wang, Fuzhong Bai, Yu Ning, Linhai Huang, and Shengqian Wang. Comparison between non-modulation four-sided and two-sided pyramid wavefront sensor. Optics express, 18(26):27534–27549, 2010.

[45] Victoria Hutterer, Ronny Ramlau, and Iuliia Shatokhina. Real-time adaptive optics with pyramid wavefront sensors: part i. a theoretical analysis of the pyramid sensor model. Inverse Problems, 35(4):045007, 2019.

[46] Frits Zernike. Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. Monthly Notices of the Royal Astronomical Society, 94:377–384, 1934.

[47] James E Millerd, Neal J Brock, John B Hayes, and James C Wyant. Instantaneous phase-shift point-diffraction interferometer. In Interferometry XII: Techniques and Analysis, volume 5531, pages 264–272. Spie, 2004.

[48] Mamadou N ' Diaye, K Dohlen, T Fusco, and B Paul. Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a Zernike phase-mask sensor. Astronomy & Astrophysics, 555:A94, 2013.

[49] M N ' Diaye, Arthur Vigan, K Dohlen, J-F Sauvage, A Caillat, A Costille, JHV Girard, J-L Beuzit, T Fusco, P Blanchard, et al. Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a Zernike phase-mask sensor-II. concept validation with ZELDA on VLT/SPHERE. Astronomy & Astrophysics, 592:A79, 2016.

[50] David S Doelman, Fedde Fagginger Auer, Michael J Escuti, and Frans Snik. Simultaneous phase and amplitude aberration sensing with a liquid-crystal vector-Zernike phase mask. Optics letters, 44(1):17–20, 2019.

[51] Ryo Tsukui, Masaru Kino, Kodai Yamamoto, and Mikio Kurita. Point-diffraction interferometer wavefront sensor with birefringent crystal. Applied Optics, 59(27):8370–8379, Sep 2020.

[52] Carolyn R Mercer and Katherine Creath. Liquid-crystal point-diffraction interferometer for wave-front measurements. Applied Optics, 35(10):1633–1642, 1996.

[53] Hiroaki Imada, Taro Matsuo, Kodai Yamamoto, and Masaru Kino. Real-time point-diffraction interferometer and its analytical formulation. Applied Optics, 54(26):7870–7878, 2015.

[54] Adrien Borne, Patricia Segonds, Benoit Boulanger, Corinne Félix, and Jérôme Debray. Refractive indices, phase-matching directions and third order nonlinear coefficients of rutile TiO 2 from third harmonic generation. Optical Materials Express, 2(12):1797–1802, 2012.

[55] NEOTRON CO.LTD. YVO4 (Yttrium Vanadate). `http://www.neotron.co.jp/crystal/7/YVO4.html`. Accessed on 2020-04-20.

[56] Mikhail N. Polyanskiy. Refractive index database. `https://refractiveindex.info`. Accessed on 2020-04-20.

[57] Remi Soummer, Laurent Pueyo, Anand Sivaramakrishnan, and Robert J Vanderbei. Fast computation of Lyot-style coronagraph propagation. Optics Express, 15(24):15935–15951, 2007.

[58] Ryo Tsukui, Masaru Kino, Kodai Yamamoto, Mina Sato, Mie Tohnishi, Akihiro Matsutani, and Mikio Kurita. Laboratory demonstration of the birefringent point-diffraction interferometer wavefront sensor. Optics Continuum, in press.

[59] Akihiro Matsutani, Mikiro Hayashi, Yasushi Morii, Kunio Nishioka, Toshihiro Isobe, Akira Nakajima, and Sachiko Matsushita. SF6-based deep reactive ion etching of (001) rutile TiO2 substrate for photonic crystal structure with wide complete photonic band gap. Japanese Journal of Applied Physics, 51(9R):098002, 2012.

[60] Lewis C Roberts Jr, Marshall D Perrin, Franck Marchis, Anand Sivaramakrishnan, Russell B Makidon, Julian C Christou, Bruce A Macintosh, Lisa A Poyneer, Marcos A van Dam, and Mitchell Troy. Is that really your Strehl ratio? In Advancements in Adaptive Optics, volume 5490, pages 504–515. SPIE, 2004.

[61] Benjamin L Gerard, Javier Perez-Soto, Vincent Chambouleyron, Maaike AM van Kooten, Daren Dillon, Sylvain Cetre, Rebecca Jensen-Clem, Qiang Fu, Hadi Amata, and Wolfgang Heidrich. Various wavefront sensing and control developments on the Santa Cruz extreme AO laboratory (SEAL) testbed. In Adaptive Optics Systems VIII, volume 12185, pages 778–790. SPIE, 2022.

[62] 国立天文台 TMT 科学諮問委員会. TMT 次期装置実現に向けたロードマップ, 2022. https://tmt.nao.ac.jp/researchers/jsac/TMT_inst_roadmap_FY2022.pdf. Accessed on 2022-12-15.

[63] 田村元秀. 新天文学ライブラリー 1 太陽系外惑星, 2015.

Appendix A

Simulation: linearity in practical cases

The linearity of the b-PDI, which was tested in Section 8.1.6, was also simulated. The simulation assumed a single wavelength of 800 nm and used the numerical model (Section 6.1.1), which reflected the actual Pinhole A (#03-40) geometry shown in Table 7.1. Assuming the input phase aberration expressed by the Zernike polynomial Z_{20}, the simulation generated the interferograms without any noise and reconstructed the phase δ_{rec}. Then, the RMS values RMS_{input} of the input phase and RMS_{output} of the reconstructed phase were compared in the same way as Section 8.1.6.

Figure A.1 shows the simulated result; the iterative algorithm enabled the linear response within $RMS_{input} \lesssim 0.7$ rad (P-V $\leq 1\lambda_C$).

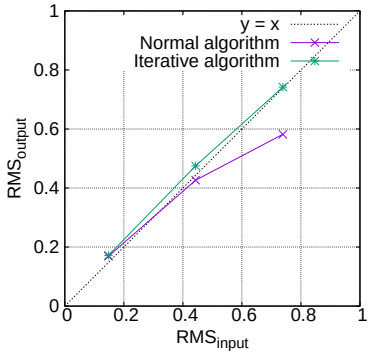


Figure A.1: Simulated response of RMS_{output} against RMS_{input} with a single wavelength of 800 nm.

Appendix B

Estimation of the number of photons

This chapter estimates the number of photons collected by a subaperture of a Tweeter WFS, assuming the parameters in Table B.1. The flux density $f(m)$ [erg sec^{-1} cm^{-2} μm^{-1}] from a m-th magnitude guide star is given by:

$$f(m) = 10^{a-0.4m}. \tag{B.1}$$

Here, a is a correction term; $a = -4.947$ at Ic-band (calculated with the values in Tamura [63]). At the loop speed F_{AO}, the Tweeter WFS's integration time is $1/F_{AO}$. In the integration time, the number of photons $N_{ph}(m)$ collected by a subaperture is expressed as:

$$N_{ph}(m) = \frac{f(m)\ L^2\ \Delta\lambda\ \lambda_C\ P}{F_{AO}\ h\ c}. \tag{B.2}$$

Here, h is Plack's constant and c is light speed. Figure B.1 shows the values of $N_{ph}(m)$ at various values of F_{AO}.

Table B.1: Parameters used in the estimation.

Parameter	Notation	Value
Central wavelength	λ_C	0.8 μm
Band width	$\Delta\lambda$	0.2 μm
Size of the subaperture	L	14.6 cm (= 380/26)
Throughput of the optics before Tweeter WFS	P	0.24

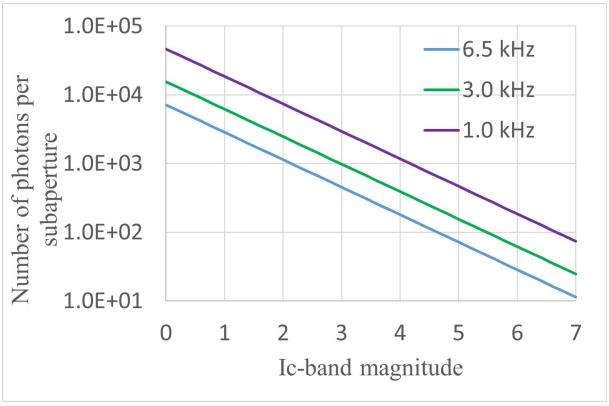


Figure B.1: The number of photons per subaperture $N_{ph}(m)$ plotted against the Ic-band magnitude m of a guide star. The assumed loop speed F_{AO} is 1.0 kHz (purple), 3.0 kHz (green), and 6.5 kHz (blue).