A High-precision Technique to Correct for Residual Atmospheric Dispersion in High-contrast Imaging Systems

P. Pathak1, O. Guyon1,2,3, N. Jovanovic1,4, J. Lozi1, F. Martinache5, Y. Minowa1, T. Kudo1, H. Takami6, Y. Hayano6, and N. Narita1

© 2016. The Astronomical Society of the Pacific. All rights reserved. Printed in the U.S.A.

Abstract

Direct detection and spectroscopy of exoplanets requires high-contrast imaging. For habitable exoplanets in particular, located at a small angular separation from the host star, it is crucial to employ small inner working angle (IWA) coronagraphs that efficiently suppress starlight. These coronagraphs, in turn, require careful control of the wavefront that directly impacts their performance. For ground-based telescopes, atmospheric refraction is also an important factor, since it results in a smearing of the point-spread function (PSF), that can no longer be efficiently suppressed by the coronagraph. Traditionally, atmospheric refraction is compensated for by an atmospheric dispersion compensator (ADC). ADC control relies on an a priori model of the atmosphere whose parameters are solely based on the pointing of the telescope, which can result in imperfect compensation. For a high-contrast instrument like the Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) system, which employs very small IWA coronagraphs, refraction-induced smearing of the PSF has to be less than 1 mas in the science band for optimum performance. In this paper, we present the first on-sky measurement and correction of residual atmospheric dispersion. Atmospheric dispersion is measured from the science image directly, using an adaptive grid of artificially introduced speckles as a diagnostic to feedback to the telescope’s ADC. With our current setup, we were able to reduce the initial residual atmospheric dispersion from 18.8 mas to 4.2 in broadband light (y- to H-band) and to 1.4 mas in the H-band only. This work is particularly relevant to the upcoming extremely large telescopes (ELTs) that will require fine control of their ADC to reach their full high-contrast imaging potential.

Key words: instrumentation: adaptive optics – atmospheric effects – planets and satellites: detection

Online material: color figures

1. Introduction

Several thousand exoplanets have thus far been discovered using indirect methods, such as transit and radial velocity, but very few using direct imaging.9 To answer questions about habitability of exoplanets, it is essential to utilize direct detection methods in order to be able to conduct spectroscopic studies. With advances in technology, direct detection of exoplanets is becoming more robust and mature. High-contrast instruments like the Gemini Planet Imager (GPI; Macintosh et al. 2014) and the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE; Beuzit et al. 2008) are now able to perform surveys and have already detected new planets (Macintosh et al. 2015; Wagner et al. 2016).

The basic architecture of a high-contrast instrument employs extreme adaptive optics (ExAO) to compensate for atmospheric turbulence and coronagraphs to suppress the light coming from the host star. A potentially significant source of coronagraphic leakage comes from atmospheric refraction once the turbulence has been accounted for, which results in an elongated point-spread function (PSF) that limits the achievable angular resolution and coronagraphic contrast. This effect can be particularly devastating if the coronagraph operates at a small inner working angle (IWA).
The PSF elongation for previous generation high-contrast instruments such as the High Contrast Instrument for the Subaru Next Generation Adaptive Optics (HiCIAO; Hodapp et al. 2008) was specified to be 1/30th of the diffraction limit ($1\lambda/D = 40$ mas in the H-band; residual dispersion goal was 1.33 mas): our experiments indicate that this was not achieved in practice (the measured residual dispersion was $\sim 12$ mas in the H-band). The latest high-contrast instrument, the Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) system (Jovanovic et al. 2015b), which is constantly evolving with time by employing the latest technologies, is going a step further in the field of direct imaging. The science requirement for SCExAO is a spread in the PSF of $\sim 1/50$th of the diffraction limit, which translates to $< 1$ mas in the H-band. At this level, the system will no longer be limited by residual atmospheric dispersion; instead, the finite stellar angular size will become the dominant term contributing to stellar leakage. For example, a solar-mass star at 10 pc has an angular size of 1 mas, demonstrating that stars normally considered as unresolved are actually partially resolved and will affect performance at these levels. The correction of the residual atmospheric dispersion is done using an atmospheric dispersion compensator (ADC), which consists of two prisms with similar dispersive properties. The amount and direction of the compensation are adjusted by rotating the prisms with respect to one another and rotating the entire assembly around the axis of propagation (Allen 1997).

An ADC model basically uses the local temperature and pressure to determine the index of refraction of air around the telescope. It does this at multiple wavelengths to guess the dispersion for different angles at the telescope. To get a correction of 1 mas ($10^{-3}$ arcsec), since the atmospheric dispersion across visible and IR is about 1 arcsec, one needs to measure the local temperature and pressure to at least a $10^{-3}$ precision. In practice, this precision is never achieved. So although the model includes these parameters precisely, the inputs may not achieve the right precision. One other term that complicates things is the humidity. Indeed, water vapor changes the index of the air. The quantity of water vapor is a key factor, as well as the wavelength observed. This effect cannot be taken into account easily. The problem with the current approach of correction is that it does not make any measurement of the actual dispersion and hence results in imperfect correction.

Adaptive optics (AO) systems include model-based ADCs, such as Keck AO (Wynne 1996) and Subaru’s AO188 (Egner et al. 2010). In this paper, we present the first on-sky measurement and correction of residual atmospheric dispersion. Here, we demonstrate that by using a grid of speckles generated by the deformable mirror (DM) of the SCExAO instrument; we can accurately measure the residual atmospheric dispersion and subsequently correct it to $\approx 1.4$ mas in the H-band. In Sections 2 and 3, we show the concept and its implementation using simulation, and Section 4 presents the on-sky results.

2. Principle

2.1. Measuring Atmospheric Dispersion

Light from an astronomical object initially undergoes refraction at the space/atmosphere boundary and then continues to refract as it propagates toward the ground. This continuous refraction is due to the increasing density and temperature of the atmosphere as one moves to lower altitudes and results in an arc-like trajectory for the light rays as they traverse the atmosphere. Refraction, of course, is a function of wavelength and hence a polychromatic PSF will look elongated at the focus of a telescope if it is not pointing at zenith. To illustrate this point, Figure 1 shows an example of a simulated PSF ($\gamma$- to H-band) for a telescope pointing at zenith with no atmospheric dispersion and compares it to an image using the same conditions, only at a lower elevation of $35^\circ$, which results in strong atmospheric dispersion. Over a broad bandwidth ($> 10$ nm or so), atmospheric refraction results in an elongation of the PSF, as shown in the Figure 1(b). An ADC, relying on a model of the atmosphere effectively compensates for most of this refraction, reducing the elongation of this PSF. However, without a direct diagnostic tool, some amount of residual dispersion remains. One could attempt to diagnose the residual dispersion ($20$–$100$ $\mu$as/nm) by directly measuring the PSF elongation, but this measurement is difficult in practice at the level required by the coronagraph due to residual wavefront error ($< 1$ mas over the H-band). Instead of attempting to measure dispersion in the PSF core itself, here we suggest using the properties of diffraction features at large separation ($> 10$ $\lambda/D$) whose location is a more sensitive function of the refraction.

Figure 2 presents the rationale behind our approach. It features a schematic representation of a broadband PSF with conveniently placed symmetric off-axis diffraction features (speckles) and shows how these are affected by the presence of refraction. With no refraction (corresponding to pointing at zenith), the images at all wavelengths will radiate from a point that coincides with the PSF core (cf. Figure 2(a)). Unlike the core of the PSF, off-axis diffraction features appear spectrally dispersed, with the shortest wavelengths no longer radiate from a point coinciding with the PSF core (cf. Figure 2(b)), but from a point that we will call the radiation center, offset from the PSF core. By measuring the distance between the radiation center and the PSF core, one can directly estimate the amount of residual dispersion.
2.2. Correcting Atmospheric Dispersion

Figure 3 shows a vector-based representation of the effect of atmospheric dispersion. The axes correspond to Cartesian coordinates in the image plane. The atmospheric dispersion vector, i.e., the direction in which the PSF is elongated, is designated with the symbol \( s \) in the figure. The dispersion vectors for the two prisms of the ADC are represented by \( p_1 \) and \( p_2 \). They have equal dispersion magnitudes and add together to generate the total ADC vector \( a \), which is also dependent on orientation and is given by

\[
a = p_1 + p_2.
\]

If \( p \) is the dispersion magnitude of each prism, and \( \theta_1 \) and \( \theta_2 \) their orientation angles, then the two individual prism dispersion vectors are decomposed in the image coordinate system as

\[
\begin{align*}
p_1 &= p \cos(\theta_1) \hat{i} + p \sin(\theta_1) \hat{j}, \\
p_2 &= p \cos(\theta_2) \hat{i} + p \sin(\theta_2) \hat{j}.
\end{align*}
\]
The ADC provides a maximum dispersion when the two prisms are aligned. When one prism is rotated 180° relative to the other (anti-aligned), the ADC results in zero compensation. For more details of how an ADC operates, refer to Egner et al. (2010).

In the case of an incomplete compensation, what is left is the residual atmospheric dispersion vector \( \mathbf{s} \), as shown in Figure 3, given by

\[
\mathbf{r} = \mathbf{a} + \mathbf{s}.
\] (3)

The goal is to measure the residual dispersion and offset the ADC to minimize the dispersion in a closed-loop operation. The steps involve the calibration of the response of the ADC prisms (prism dispersion magnitude \( p \)) and the calculation of on-sky dispersion. If we assume that the ADC prism angles are known all the time, then the components in the image coordinate system can be written as

\[
\begin{aligned}
  a_x &= p \cos(\theta_1) + p \cos(\theta_2) \\
  a_y &= p \sin(\theta_1) + p \sin(\theta_2) .
\end{aligned}
\] (4)

In the same coordinate system, the residual vector can be decomposed into

\[
\begin{aligned}
  r_x &= s_x + a_x \\
  r_y &= s_y + a_y .
\end{aligned}
\] (5)

To calibrate the response of the ADC prisms, we assume that, over small timescales, the on-sky dispersion vector \( \mathbf{s} \) is constant. As shown in Figure 3, by making two measurements of residual \( \mathbf{r} \) and \( \mathbf{r}' \) using two known positions of ADC, we can eliminate the atmospheric contribution \( \mathbf{s} \). By rotating the ADC by an angle \( \theta_{\text{off}} \), the new ADC vector \( \mathbf{a}' \) is given by

\[
\begin{aligned}
  a'_x &= p \cos(\theta_1 - \theta_{\text{off}}) + p \cos(\theta_2 - \theta_{\text{off}}) \\
  a'_y &= p \sin(\theta_1 - \theta_{\text{off}}) + p \sin(\theta_2 - \theta_{\text{off}}) .
\end{aligned}
\] (6)

The new residual vector \( \mathbf{r}' \) is then decomposed into

\[
\begin{aligned}
  r'_x &= s_x + a'_x \\
  r'_y &= s_y + a'_y .
\end{aligned}
\] (7)

By subtracting Equation (7) from Equation (5), we get

\[
\begin{aligned}
  r_x - r'_x &= a_x - a'_x \\
  r_y - r'_y &= a_y - a'_y .
\end{aligned}
\] (8)

Now substituting Equations (4) and (6) into Equation (8), we get

\[
\begin{aligned}
  p \times l &= r_x - r'_x \\
  p \times m &= r_y - r'_y .
\end{aligned}
\] (9)

where \( l \) and \( m \) are given by

\[
\begin{aligned}
  l &= \cos(\theta_1) + \cos(\theta_2) - \cos(\theta_1 - \theta_{\text{off}}) - \cos(\theta_2 - \theta_{\text{off}}) \\
  m &= \sin(\theta_1) + \sin(\theta_2) - \sin(\theta_1 - \theta_{\text{off}}) - \sin(\theta_2 - \theta_{\text{off}}) .
\end{aligned}
\] (10)

By solving Equation (9), we get

\[
\begin{aligned}
  p &= \left( \frac{(r_x - r'_x)^2 + (r_y - r'_y)^2}{l^2 + m^2} \right)^{1/2} .
\end{aligned}
\] (11)

Since we know the prism angles \( \theta_1 \) and \( \theta_2 \), the offset applied to the prisms \( \theta_{\text{off}} \), the value of \( l \) and \( m \), can be deduced using Equation (10). Then, by substituting the values of the measured residual dispersion \( (r_x, r_y) \), \( (r'_x, r'_y) \) and \( l \) and \( m \) into Equation (11), the magnitude of the prism dispersion vector \( p \) can be calculated.

Once \( p \) is determined from the measurements, the vector \( \mathbf{a} \) is known for any ADC angle using Equation (3). For any measurement of the residual vector \( \mathbf{r} \), we can deduce the on-sky dispersion vector \( \mathbf{s} \) using Equation (3). Finally, once we know the dispersion \( \mathbf{s} \), we can determine the new ADC angle that compensates for it.

### 3. Simulation

To test the concepts explained in the previous section, we present the result of simulations that further characterize our approach. As explained by Jovanovic et al. (2015a) and Martinache et al. (2014), artificial speckles at specific locations can deliberately be introduced in an image by a DM featuring a sufficiently large number of actuators, within a finite region called the control region of the DM. We use this convenient approach to generate symmetric pairs of speckles in the image that will be used to sense the amount of residual refraction in focal plane images.

The simulations take the Subaru telescope pupil geometry into consideration, as shown in Figure 4(a). For the images in Figures 4(b), (c), and (d), a 12.1 mas/pix plate scale was used to match that of SCExAO’s internal NIR camera. Figure 4(b) shows what a perfect monochromatic PSF looks like.

Panels (c) and (d) of Figure 4 feature simulated broadband images, corresponding to the combined y- to H-band flux. Low-amplitude sinusoidal modulations were added to the DM to produce pairs of off-axis speckles along the diagonals of \( x \) and \( y \) axes of the image. For panel (c), no differential refraction is implemented so that the spectrally dispersed off-axis speckles appear to radiate from the core of the PSF. In panel (d), 220 μas/nm of differential refraction was introduced: the radiation center of the off-axis speckles is no longer pointed toward the PSF core.
3.1. Extracting Atmospheric Dispersion

In order to estimate the amount of residual atmospheric refraction in a given image, the distance between the core of the PSF and the radiation center must be measured precisely. Even with some prior knowledge, the careful mapping of non-symmetric structures is a non-trivial task in practice. We have however empirically developed an algorithm that provides satisfactory performance that is summarized in Figure 5:

- Panel (a) shows a schematic representation of an image affected by residual refraction. We use $d_x$ and $d_y$ to refer to the horizontal and vertical offset between the PSF core and the radiation center, respectively.
- Panel (b) shows a similar representation for which the PSF core has been masked out. The initial guess is that the PSF core corresponds to the radiation center (i.e., there is no residual dispersion). The original (blue) image is then radially stretched from the location of the original PSF core by a finite amount (in red). One can observe that for a sufficiently large stretch factor (of the order of 20% in practice), the spectrally dispersed structures present in both images (the blue and the red) no longer overlap. This difference can be quantified by the norm of the difference between the two images $\mathcal{L}$.
- The next step is to shift the assumed center of radiation and re-stretch the original image to find the point at which the norm $\mathcal{L}$ is minimized (i.e., the speckles from the original image and those from the stretched image overlap most). In this case, a raster scan is conducted but other scanning algorithms could also be used.
- Panel (c) shows a final case featuring both the original (blue) and the stretched (red) version of the images, only

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Theoretical pupil and corresponding PSFs assuming no turbulence applied (i.e., off-sky). All PSF images are polychromatic. (a) Subaru Telescope pupil geometry. (b) PSF corresponding to the pupil. (c) Same PSF including four artificial speckles generated by applying a sinusoidal modulation in phase across the pupil. (d) Same PSF with speckles including the affect of atmospheric dispersion. (A color version of this figure is available in the online journal.)}
\end{figure}
this time, the latter was stretched from the correct radiation center: the speckles now have a significant overlap and our criterion of minimizing $\mathcal{L}$ is satisfied.

The result of one such computation is shown in Figure 6 for two cases, in the absence of residual refraction in panel (a) and with 18.8 $\mu$as/nm of refraction in panel (b). As expected in the first case, $\mathcal{L}$ is minimum when the PSF core and the radiation center are identical (i.e., $d_0 = 0$ and $d_y = 0$). In the second scenario, the location of the minimum of the $\mathcal{L}$ is offset. The distance between the PSF core and the radiation center is about twice this offset, i.e.,

$$r_i \simeq 2 \times d_i, \quad i \in [x, y].$$

This relationship is used to calculate the on-sky dispersion in Section 4. By this method, one can measure the residual dispersion down to sub-pixel accuracy.

### 3.2. ADC Simulation

The operation of an ADC was simulated to test the measurement and correction of the dispersion on a simulated PSF. As described in Figure 7, the first part of the simulation involves the calibration of the ADC. The inputs used in simulation are prism angles $\theta_1$ and $\theta_2$ and the magnitude of the prism dispersion vector $p$, which is an unknown constant. By using these parameters, the ADC vector $\mathbf{a}$ is calculated using the equations presented in Section 2. By using the measured vectors $\mathbf{r}$ and $\mathbf{a}$, the on-sky dispersion vector $\mathbf{s}$ can be calculated. In a second step, an offset $\theta_{\text{off}}$ is applied to $\theta_1$ and $\theta_2$, which gives the new ADC vector $\mathbf{a}'$ and residual vector $\mathbf{r}'$. By using residual vectors $\mathbf{r}$, $\mathbf{r}'$ and the offset angle $\theta_{\text{off}}$, the prism dispersion magnitude $p$ can be calculated. In the second part of the simulation using the measured residual dispersion vector $\mathbf{r}$, the prism angles and $p_{\text{calculated}}$ offset angles $\delta \theta_1$ and $\delta \theta_2$ are calculated, which will give us a better correction of dispersion.
4. Results

As explained in Jovanovic et al. (2015a) and Martinache et al. (2014), artificial speckles can be generated by diffracting the PSF core with a DM having a large number of actuators. In this case, SCExAO’s DM was used, which has a total of 2000 actuators, with 45 actuators across the pupil. The DM can be modulated to create a grating structure in the form of a sine wave. The distance between the PSF core and the resulting speckles is a function of the number of periods across the pupil. The more cycles per aperture the sine wave has, the further a speckle is projected from the PSF. With 45 actuators across the pupil, the furthest speckles can be placed is 22.5 λ/D from the PSF (Jovanovic et al. 2015b). The brightness of the speckles can be controlled by adjusting the amplitude of the sine wave. Previously, this ability to arbitrarily generate speckles has been used to systematically remove speckles in the PSF Halo (Martinache et al. 2014) and more recently for high-precision astrometry using incoherent speckles (Jovanovic et al. 2015a). For our tests, on-sky speckles were placed at 22.5 λ/D, with a 100 nm rms amplitude. Images were taken using a near-infrared (NIR) camera (320 × 256 pixel InGaAs detector). The target was Beta Leo (spectral type A3, R-mag = 2.08, H-mag = 1.92) on SCExAO’s engineering night of 2015 April 2. These data were collected from AO188, the Subaru Telescope facility AO instrument. In closed loop, it offers Strehl ratios in the H-band between 20% and 40% (Minowa et al. 2010). To capture the elongated speckles, the light from the y- to H-band was captured by the NIR camera. For image processing, an averaged dark was calculated from a cube of 1000 dark frames, then subtracted from the science images. The hot pixels were also removed. At the time of the test, the telescope elevation was 43° and no ADC correction was applied to highlight the presence of atmospheric dispersion. As shown in Figure 8, the speckles do not point to the PSF core as expected.

The on-sky correction of residual atmospheric dispersion was achieved on the target Alpha Ari (spectral type K1, R-mag = 1.15, H-mag = −0.52) on SCExAO’s engineering night of 2015 October 30, using the same specifications. The results of the on-sky measurement and correction of residual atmospheric dispersion are presented in Figure 9. Figure 9(a) shows the added speckles and the PSF core, with superimposed lines, showing that the radiation center lies away from the PSF core. Figure 9(b) shows the added speckles after correcting for residual dispersion. Here, the intersection point lies closer to the PSF core. After determining the residual dispersion and the appropriate angles of the ADC that would minimize it, the ADC was driven to a new position to make the radiation center coincide with the PSF core. The method outlined in Section 3.1 was used to find the radiation center, and the results are presented in Table 1. The values of dx and dy shown in the table represent the distance between the radiation center and the PSF core, in units of mas converted from pixel values (using the known scaling factor of 12.1 mas/pix in our NIR camera). The
values of $d_x$ and $d_y$ are then converted in PSF elongation by the relationship between distance and dispersion given by Equation (12). Table 1 shows that in a single iteration, we have reduced the residual atmospheric dispersion from $20.9 \mu\text{as}/\mu\text{m}$ to $4.7 \mu\text{as}/\mu\text{m}$, which corresponds to 4.2 mas in the y- to H-band and 1.4 mas in the H-band alone, which is close to our science requirement of 1 mas in the H-band.
5. Discussion

The uncertainty in the values of $dx$ and $dy$ given in Table 1 were determined by calculating the standard deviation of those two values from a cube of 1000 images. This term consists of a combination of errors including the determination of the PSF core to sub-pixel accuracy as well as locating the radiation center that will be affected by the stability of the speckles due to both residual turbulence and possible blurring during an exposure, and the chromatic component to tip-tilt. A proper error budget is beyond the scope of this work, but with careful consideration, we believe the residual dispersion could be further optimized to the desired value of <1 mas.

One approach that could potentially improve the correction could be to exploit incoherent artificial speckles rather than the coherent versions used in this body of work. As explained in Jovanovic et al. (2015a), incoherent speckles can be created by rapidly modulating the phase of the artificial speckle grid by $\pi$ during a single exposure. In this way, speckles that do not interact with the underlying speckle halo can be created. Using these speckles, the stability of their brightness and centroid can be improved, which would improve the determination of the radiation center.

Finally, by using a method that uses the final science image to correct for residual atmospheric dispersion, it is possible to correct for dispersion resulting from optics internal to the instrument. This term is not taken into account in ADC models, and so direct measurement of this effect in the final focal plane is the only solution.

6. Conclusion

In this work, we demonstrated the first on-sky measurement and correction of residual atmospheric dispersion using an adaptive speckle grid. This technique can be used by other AO systems, which do not employ high actuator DMs to create speckle grids, by using a diffractive grid in the pupil. This concept will be extended to work in conjunction with coronagraphs, where the location of the PSF behind the coronagraph can be found by cross-correlating the speckles in the grid. In this work, we have canceled one of the leading noise terms which prevents direct imaging of exoplanets. This work will be valuable in the field of ground-based high-contrast imaging of habitable exoplanets in the era of extremely large telescopes (ELTs) because it shows the path to correct atmospheric dispersion to the highest degree compared to traditional approaches. It can also be incorporated in atmospheric dispersion corrector design for ELTs (Bahrami & Goncharov 2011).

The authors acknowledge support from the JSPS (Grant-in-Aid for Research 23340051 and 26220704).

*Facility: Subaru Telescope.*

References

Allen, C. W. 1997, Astrophysical Quantities (3rd ed.; London: The Athlone Press)
Bahrami, M., & Goncharov, A. V. 2011, *OE Expr*, 19, 17099
Beuzit, J.-L., Feldt, M., Dohlen, K., et al. 2008, in Proc. SPIE 701418
Egner, S., Yuji, I., Watanabe, M., et al. 2010, *Proc. SPIE*, 7736, 77364V
Hodapp, K. W., Suzuki, R., Tamura, M., et al. 2008, *Proc. SPIE*, 7014, 701419
Jovanovic, N., Guyon, O., Martinache, F., et al. 2015a, *ApJ*, 813, 24
Jovanovic, N., Martinache, F., Guyon, O., et al. 2015b, *PASP*, 127, 890
Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Sci*, 350, 64
Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, *PNAS*, 111, 12661
Martinache, F., Guyon, O., Jovanovic, N., et al. 2014, *PASP*, 126, 565
Minowa, Y., Hayano, Y., Oya, S., et al. 2010, *Proc. SPIE*, 7736, 77363N
Wagner, K., Apai, D., Kasper, M., et al. 2016, *Sci*, 353, 673
Wynne, C. G. 1996, *MNRAS*, 282, 863