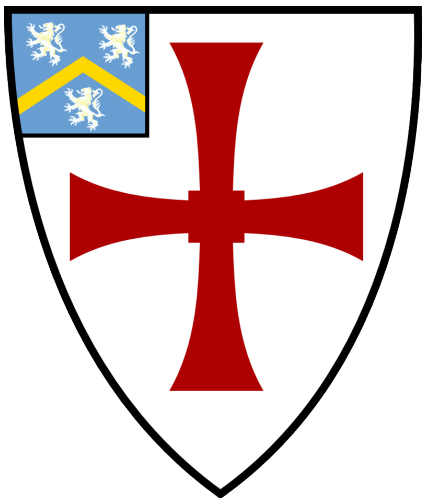
Modelling of Long Exposure Wavefront Errors for Extremely Large Telescopes

Polly Gill

A thesis presented for the degree of

Doctor of Philosophy



Centre for Advanced Instrumentation

The University of Durham

United Kingdom

December 2023

Modelling of Long Exposure Wavefront Errors for Extremely Large Telescopes

Polly Gill

Abstract

Quasi-static aberrations are wavefront distortions that vary on a timescale longer than that typically controlled by the active optics control loops used to maintain the image quality within a large telescope. The source of these errors can be due to gravitational or thermal flexure of the telescope; or poorly averaged atmospheric turbulence residuals.

Following the method of Gordon et al. (2011), a simulation was used to explore the averaging rate of atmospheric turbulence residuals for a phase screen translating across the telescope aperture. Crucially, the rate at which the individual modes average out is dependent upon wind direction, meaning that both the C_n^2 and wind velocity profiles can have a large effect on the modal variance of observed quasi-static aberrations.

We have verified that the use of a Zernike covariance matrix library is both viable and useful when considering a multi-layered atmospheric profile with variable velocity dispersions. Using this library we tested the effects of finite averaging time in an active optics model, calculating the residual error generated by a correction from an off-axis guide star. We have used this model for varying atmospheric velocity dispersions and exposure times, for both existing telescopes and ELT scales.

Supervisors: Tim Butterley and Tim Morris

Acknowledgements

I need to first thank my brilliant supervisory team of Tims, Butterley and Morris, for all of their help and support through not only a PhD but also a global pandemic. Thank you for being so very generous with your patience and your time.

Also to my office-mates past, present and honorary: Chris, Josh, Aurélie, Lily, Craig, Jeremi and Katy; thank you for all of your encouragement, particularly in the last few months of hard drive panic. To anyone who has opened our door hopefully shaking a coffee mug, thank you for both the distraction and the caffiene.

Of course an enormous thanks must also go to my family. Mum, Dad and George, I'm so grateful for all of your support. Peter, I could not have done this without you, thank you so much.

Finally, I dedicate this thesis to my wonderful Grandad. You could draw a straight line across my life between your genuine care for my wobbly maths skills some nineteen years ago, and the science I adore now. I cannot thank you enough.

Contents

Declaration

The work in this thesis is based on research carried out at the Centre for Advanced Instrumentation, Department of Physics, University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification, and it is the sole work of the author unless referenced to the contrary in the text.

Some of the work presented in this thesis has been published in journals and conference proceedings - the relevant publications are listed below.

Publications

Gill, Polly, Tim Butterley, and Tim Morris. "Effects of wind velocity profiles on turbulence-induced quasi-static aberrations." Adaptive Optics Systems VIII. Vol. 12185. SPIE, 2022.

List of Figures

List of Tables

Nomenclature

AcO Active Optics

AO Adaptive Optics

ANGuS All Natural GUide Star

CRCM Coordinate Rotation of the Covariance Matrix

DM Deformable Mirror

FT Fourier Transform

FFT Fast Fourier Transform

FoV Field of View

FWHM Full Width Half Maximum

NGS Natural Guide Star

NOT Nordic Optical Telescope

PSD Power Spectral Density

PSF Point Spread Function

RMS Root Mean Square

RMSE Root Mean Square Error

SCAO Single-Conjugated Adaptive Optics

SCIDAR Stereo-SCIntillation Detection And Ranging

SLODAR SLOpe Detection And Ranging

VISTA Visible and Infrared Survey Telescope for Astronomy

VLT Very Large Telescope

WFE Wave-front Error

WFS Wave-front Sensor

Chapter 1

Introduction

Evidence of tracking star patterns can be traced all the way back to the Assyro-Babylonians in the twelfth century BCE. "Three Stars Each" is suggested to be the earliest catalogue, and is considered a calendar of sorts, containing three stars for each month. For the vast majority of astronomy, we looked up at the night sky with no tools to aid us. The advent of the telescope is muddy, but the first reflecting telescope is generally attributed to Newton in the 17th century and was just 2 inches in diameter (Newton, 2014)[†]. After this, technological advancements began to push telescopes wider to gather more light - in just over 100 years Herschel had developed a reflecting telescope 1.2m in diameter at the primary mirror. As time passed the techniques for mirror creation and refinement improved, meaning existing telescopes can be up to 10m. The next generation of planned telescopes, such as the ESO-ELT will be 39m in diameter - over 700 times wider than Newton's first attempt 355 years ago. As the telescope sizes get larger, correcting their optical errors becomes both more necessary and complicated.

An example telescope layout is shown in figure 1.1. Any distortion or misalignment of the optical elements (shown in blue) will cause aberrations of the telescope image and worsen the imaging performance (Noethe, 2002). If these aberrations can be measured independently then corrections can be applied in real time by adjusting the telescope mirrors. This computing control system is shown in purple in figure

[†]The letter was written in 1672 and digitally uploaded by the Royal Society in 2014.

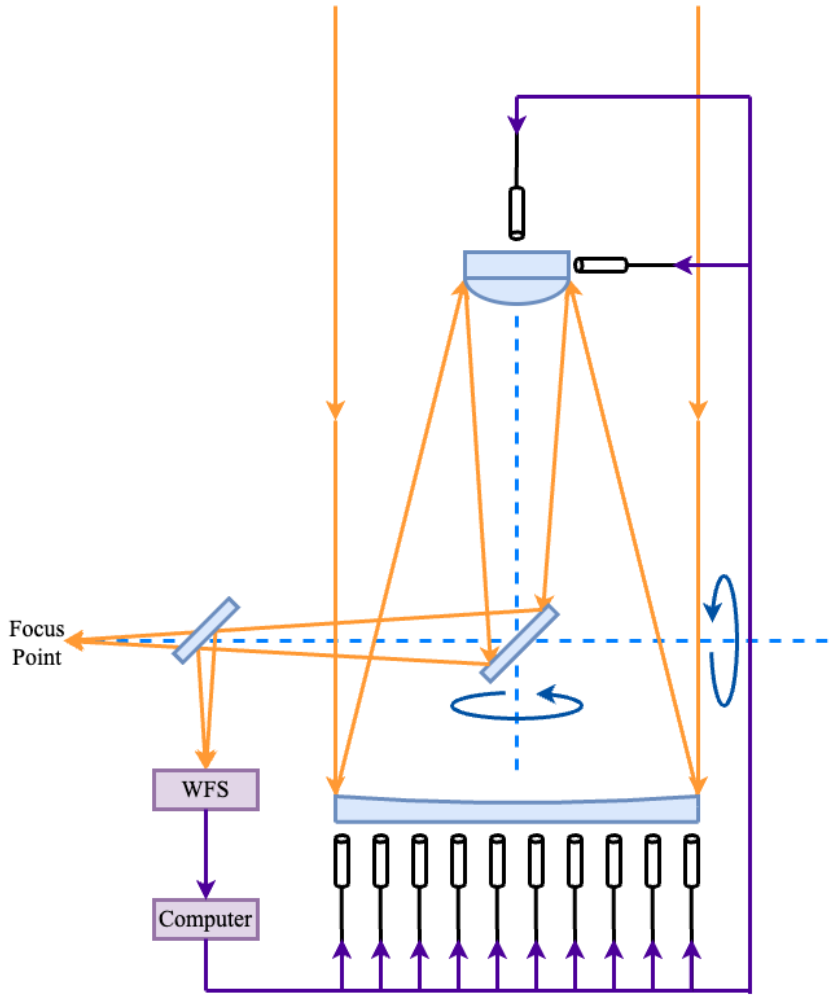


Figure 1.1: Diagram showing a simple active optics setup. Light is reflected through a telescope and before it reaches the focus point, some is split off into a WFS which reconstructs the wavefront. A computer system is used to control the shape and positions of the mirrors, correcting for the aberrations caused by the incoming wavefront, the telescope itself, or both. If this process is successful the image at the detector placed at the focus point will be stationary and diffraction limited.

1.1.

If the large primary mirror is one monolithic piece of glass then its shape can be changed with the use of actuators which "poke" the mirror into a a different shape, to improve the performance of the telescope - as seen on the William Herschel telescope, or the VLT. Mirrors with this variable shape are known as active mirrors. The active mirror supports for the VLT are shown in figure 1.2. As the telescope

diameter gets larger it becomes unfeasible to use a continuous primary mirror. Instead it can be split into segments, with a series of smaller mirrors acting as one large diameter. With this setup a telescope control system can change the position and tilt of the different segments to account for aberrations to the wavefront. This design will be used in the ESO-ELT, where 798 hexagonal mirrors with individual diameters of 1.4m will be put together to form a 39m diameter primary mirror.

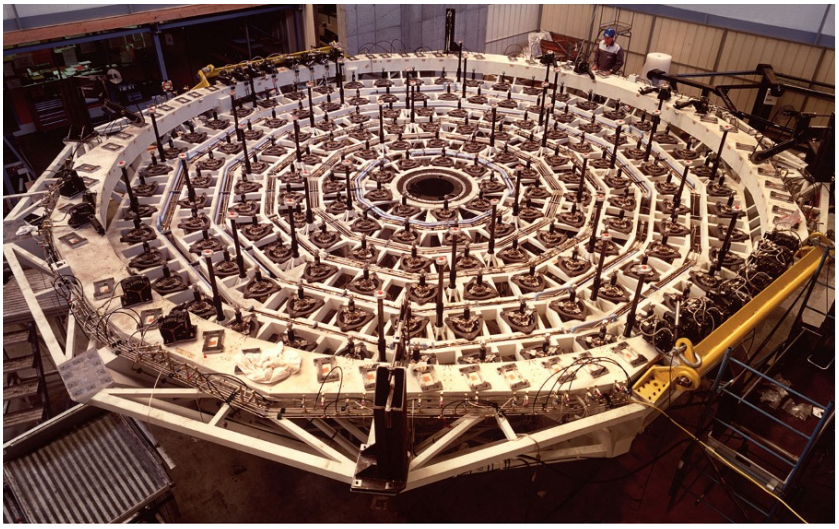


Figure 1.2: The 150 computer-controlled supports for the VLT primary mirror. Credit: ESO, PR Image eso9940a

An AcO system measures the slow moving errors inherent to the telescope itself, allowing for this to be corrected for in real time. For example as the telescope pointing and tracking changes, the gravitational vector will shift, causing the optical components to sag and continually change shape on a slow time scale. Flexure can also be caused by changes in temperature warping the mirror materials. Wind buffeting around the dome may also changes to the telescope alignment. The use of such a correction system was first demonstrated on sky by the New Technology Telescope (Wilson et al., 1991).

An AcO system updates on the timescale of tens of seconds. At higher frequencies there is still an atmospheric error that cannot be compensated for (Guisard et al., 2000). High frequency atmospheric aberrations evolve faster than the active optics can account for, so act as a noise source. Therefore AcO is now often used in

tandem with an adaptive optics system, which accounts for the quickly evolving aberrations caused by the atmospheric turbulence. The turbulence changes the path that star light takes through our atmosphere, which inhibits a telescope's ability to focus accurately and degrades its image quality (Wilson et al., 1987). Adaptive Optics (AO) was first proposed in the mid 20^{th} century in Babcock (1953), the optics within the telescope must adapt with the moving atmosphere. An AO system updates on the timescale of milliseconds in order to correct for the ever evolving atmospheric effects - much faster than AcO, but the corrections are also much smaller in magnitude.

Given that the atmospheric turbulence varies across the sky, the correctable area from a single guide star is limited. As the separation between the guide star and the science target increases the turbulence in each line of sight becomes more dissimilar. This is referred to as anisoplanatism (Fried, 1982). Whilst VISTA has a smaller diameter than the VLT, it covers a wider field of view as a survey telescope, and so needs three off axis guide stars (Terrett et al., 2004). At the VLT they attempt to remove the remainder fast moving errors with a filter (Guisard et al., 2000). Next generation 20m+ telescopes will require AcO to function, Clénet, Y. et al. (2010), but their large diameter will require multiple guide stars. The ESO-ELT will be equipped with six laser guide stars, and will use up to five natural guide stars in addition (Rousset et al., 2010). The full adaptive optics requirements call for the ELT to have multiple active mirrors, meaning unaccounted for errors that are fed back into the control system of the ELT may cause focal plane distortions (Rodeghiero et al., 2021).

Gordon et al. (2011) showed that the effects of atmospheric aberrations should average out towards zero with long exposure times, which is correlated to telescope diameter: $\tau \propto D^{8/3}$. This will leave only the instrument error for the active optics system to correct. These long averaging times could become an issue for the performance of next generation ELT telescopes and their simulation is the main focus of this thesis.

In addition to purely considering the long exposure times necessitated by these extremely large diameters, we have also explored the errors between multiple lines of sight. If using an off-axis guide star, the atmospheric aberrations along both lines of sight will be slightly different. This will cause an error in the correction, purely caused by the atmosphere. How this scales with exposure time could affect the active optics performance. Assuming Taylor's frozen flow, the exposure time is intrinsically linked to the wind speed. Throughout this thesis we test how the atmospheric residual error scales with telescope diameter, wind velocity and separation between a guide star and the on-axis target, focusing on next generation ELT scale telescopes.

1.1 Synopsis

Chapter 2 discusses the necessary background turbulence and simulation theories, including the use of phase screens for a multi-layer atmospheric profile and our initial long exposure method.

In chapter 3 we explore how the variance and RMS for coefficients from individual Zernike modes are affected by changes to both wind speed and wind direction. We start with a simple atmosphere consisting of one turbulent layer and then extend to a multi-layered atmospheric model, with layers moving at different velocities. This uses both theoretical, helically varying angular dispersions and based on on-sky data from both SCIDAR and the Paranal ambient atmospheric conditions database. We show how this changes the time necessary for a single line of sight to average to an acceptable threshold of error.

Extending our research into the effects of different wind directions on two lines of sight could become computationally intensive and time consuming, particularly if using a multi-layered atmospheric model. In chapter 4 we describe the construction of a more efficient covariance library method and justify the parameter spaces we choose.

We utilise this covariance library in chapter 5 to create an active optics model, calculating the residual error between one off axis guide star and the on axis science target. We prove this model against a tested independent Monte Carlo simulation ANGuS. This is shown for changes to: the angular separation between the lines of sight, the exposure time used to measure the reference guide star and the wind velocity of the atmospheric profile. We demonstrate how the residual error also scales with telescope diameter and the possible advantages of varying modal averaging times. The final chapter, chapter 6 contains our conclusions and possible future work.

Chapter 2

Methods

2.1 Turbulence Theory

Sunlight causes large temperature variations in the Earth's atmosphere, causing pressure differentials which create wind. Initially the air moves in laminar type flow with near uniform changes in velocity (Taylor, 1938). In certain conditions, the flow becomes turbulent. This change occurs when the Reynolds number is large enough - the exact value varies between different mediums (Roddier, 1981). The Reynolds number depends on the velocity V_0 and length L of the flow and kinematic viscosity ν_0 of the fluid as

$$Re = \frac{V_0 L}{\nu_0}. \tag{2.1}$$

Pockets of air form, which vary in size and temperature. These pockets, or turbulent eddies, differ in density and therefore refractive index. The eddies alter the path of light travelling through our atmosphere and degrade potential image quality of ground based astronomical telescopes. One solution to this would be to place telescopes above the atmosphere; but space based telescopes are limited in size and therefore are capable of gathering less light and have a larger diffraction limit (Énard et al., 1996). Alternatively, if a telescope could remove the atmospheric aberrations, larger ground based telescopes would still be viable. In order to mitigate

the atmospheric effects, the turbulence and movement of air in the atmosphere has to be characterised and a statistical model which accounts for random nature of the turbulence found (Roddier, 1999).

Light from a star or other distant point source travelling through our atmosphere can be considered a plane wave, meaning the amplitude and phase of the electric field are constant with respect to position. The width of a point source image through a circular aperture with diameter D is diffraction limited, with Airy disk radius of $1.22\lambda/D$ (Hecht, 2012). However, if that wave passes through the turbulent atmosphere, the eddies act like localised positive and negative lenses, changing the curvature of the wavefront.

Energy cascade theory dictates that the kinetic energy of turbulent flow moves from large scale eddies at the size of the outer scale L_0 down into smaller pockets (Von Karman (1948), Kolmogorov (1991)). When the eddies are small enough the energy is dissipated through viscous friction between air molecules. The scale of these small eddies is referred to as the inner scale l_0. Between these extremes is the inertial subrange. The inner scale varies from a few millimetres when near the ground to a few centimetres higher up (Roddier, 1981). A larger inner scale limits the turbulent flow, weakening the turbulence, but has minimal effects on atmospheric propagation (Valley, 1979). The outer scale can reach hundreds of metres high in the atmosphere, and tends to the height above the surface when close to it (Roddier, 1981).

Modelling the effects of turbulence requires a statistical approach as opposed to trying to describe the variation of refractive index across space and time in a closed, analytical solution. By assuming that the eddies in the inertial sub-range are homogeneous and isotropic in small sections, characteristics like velocity and refractive index will change in static increments. The outer scale is the limit at which the turbulence can no longer be considered isotropic.

Kolmogorov studied the difference in mean square velocity between two points

(Kolmogorov (1991) translation of original paper) separated by vector $\mathbf{r}$. The structure tensor can be expressed in terms of the velocities v_i and v_j as:

$$D_{ij} = \langle [v_i(\mathbf{r_1} + \mathbf{r}) - v_i(\mathbf{r_1})][v_j(\mathbf{r_1} + \mathbf{r}) - v_j(\mathbf{r_1})] \rangle, \tag{2.2}$$

for Kolmogorov turbulence. This can be simplified by making a few assumptions about the atmosphere. If the atmosphere is locally homogeneous then the velocity only depends on the vector $\mathbf{r}$ and if the atmosphere is locally isotropic then the velocity only depends on the magnitude of $\mathbf{r}$. In addition if the turbulence is incompressible then equation 2.2 can be simplified to a structure function:

$$D_v = \langle [v_r(\mathbf{r_1} + \mathbf{r}) - v_r(\mathbf{r_1})]^2 \rangle \tag{2.3}$$

If the separation is small enough to fit within the inertial sub-range, but larger than l_0, then this can be re-expressed as:

$$D_v = C_v^2 r^{2/3} \tag{2.4}$$

Where C_v^2 is the velocity structure constant - a measure of the energy within the turbulence. The turbulence can also be described with a refractive index structure function.

$$D_n(r) = C_n^2 r^{2/3} \text{ for } l_0 \ll r \ll L_0 \tag{2.5}$$

where C_n^2 is the refractive index structure constant which has units $m^{-2/3}$. A larger C_n^2 value indicates larger amplitude aberrations caused by the atmosphere. Assuming the atmosphere has a mean index of refraction $\langle n(\mathbf{r}) \rangle$, where $\langle\ \rangle$ indicates the ensemble average, and a varying part $n_1(\mathbf{r})$ gives the covariance of the refractive index field as:

$$B_n = \langle n_1(\mathbf{r_1} + \mathbf{r}) n_1(\mathbf{r_1}) \rangle \tag{2.6}$$

From here the PSD of the variations in refractive index can be determined by the Fourier transform of equation 2.6:

$$\Phi_n(\kappa) = \frac{1}{(2\pi)^3} \int d^3\mathbf{r} B_n(\mathbf{r}) e^{-i\mathbf{K}.\mathbf{r}} \tag{2.7}$$

where $\mathbf{K}$ is the 3D spatial wave number. By switching to spherical coordinates where $\mathbf{K} = (K, \Theta, \phi)$ and using equation 2.5 reduces equation 2.7 to:

$$\Phi_n(\kappa) = \frac{5}{18\pi} \int_{l_0}^{L_0} dr \sin(Kr) r^{-1/3} \tag{2.8}$$

If the integral limits diverge to $l_0 \to 0$ and $L_0 \to \infty$, this now represents the Kolmogorov spectrum:

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3} \tag{2.9}$$

where $\kappa = 2\pi(f_x \mathbf{i} + f_y \mathbf{j})$ is the spatial frequency Kolmogorov (1991). The Von Karman PSD follows in much the same way but includes both the inner and outer scale:

$$\Phi_n(\kappa) = 0.033 C_n^2 \frac{exp(\kappa^2/\kappa_m^2)}{(\kappa^2 + \kappa_0^2)^{11/6}} \text{for } 0 \ll \kappa \ll \tfrac{1}{l_0} \tag{2.10}$$

where $\kappa_0 = 2\pi/L_0$ and $\kappa_m = 5.92/l_0$ (Von Karman, 1948). In simulation, we assume that different layers of turbulence are separated vertically enough that they can be treated as statistically independent layers with their own strengths $C_n^2(h)$, meaning each layer can be represented independently by a different phase screen.

The Fried parameter r_0 or Fried coherence length, is a measure of the total strength of the turbulence integrated along a line of sight:

$$r_0 = [0.423 k^2 (\cos\gamma)^{-1} \int dh C_n^2(h)]^{-3/5} \tag{2.11}$$

where γ is the zenith angle and $k = 2\pi/\lambda$ (Fried, 1966). As can be seen from Eq. 2.11, $r_0 \propto \lambda^{6/5}$ meaning longer wavelengths have a larger Fried parameter and therefore the turbulence has less effect (Roddier, 1981). For optical wavelengths, r_0 typically has magnitude of tens of centimetres at astronomy sites. The phase power spectrum for Kolmogorov turbulence can be expressed in terms of r_0 as:

$$\Phi(\kappa) = 0.023 r_0^{-5/3} \kappa^{-11/3} \tag{2.12}$$

From here the phase structure function can be calculated for Kolmogorov turbulence (Rao et al., 2000):

$$D_\Phi(r) = 6.88 \left(\frac{|r|}{r_0}\right)^{5/3}. \tag{2.13}$$

2.1.0.1 Temporal Effects

It can be assumed that light travels so quickly that it will travel through an eddy before the properties of the eddy has time to change. Refractive index can therefore be thought of as independent of time over very short time scales. To allow for changes in time, Taylor's frozen-flow turbulence hypothesis treats the eddies as fixed and are all moved together by a single wind velocity as it crosses the pupil (Taylor, 1938).

The coherence time τ_0 represents the timescale over which the turbulent aberrations are correlated as

$$\tau_0 = 0.314\frac{r_0}{\vec{v}}, \tag{2.14}$$

where $\vec{v}$ is the effective wind speed:

$$\vec{v} = \left(\frac{\int_0^\infty C_n^2(h)V(h)^{5/3}dh}{\int_0^\infty C_n^2(h)dh}\right). \tag{2.15}$$

where $V(h)$ is the wind speed as a function of altitude. τ_0 defines time the adaptive optics system has to correct the turbulence before it changes significantly. By definition then, τ_0 represents the the minimum update rate of an adaptive optics system in order for it to correct the same turbulence the telescopes sees (Roddier, 1981).

The coherence time is dependent on both the strength and the velocity of the turbulence.

If the exposure time is short, light passing through the atmospheric turbulence will produce a distorted image. If observing with a larger telescope, an image of the star may break into speckles. Imaging on longer timescales allows the turbulence to randomly evolve, producing a blurred image, referred to as seeing.

For long exposures, the image of a point source will change from speckles of size λ/D to a $\sim$ Gaussian Point Spread Function (PSF) with a Full Width Half Maximum (FWHM) of λ/r_0 as it is now limited by the turbulence, rather than the telescope size (Fried, 1966). Longer wavelengths will have a larger Fried parameter, so diffraction limited images are easier to achieve in the infrared than in the optical.

2.1.1 Simulating Turbulence Using Phase Screens

To simulate the effects of the atmosphere on light travelling through it, phase screens are used to represent discrete thin layers of turbulence which can be translated across a telescope pupil to represent frozen flow. Multiple phase screens can be generated to represent more than one layer of turbulence. The method presented here to generate a phase screen is taken from Ellerbroek and Cochran (2002). To generate a phase screen, consider the effects a thin layer of turbulence will have on a wavefront. We can express a complex wavefront U of a single wavelength λ, with amplitude A and phase ϕ as:

$$U(x,y) = A(x,y)exp(i\phi(x,y)) \tag{2.16}$$

Variations in the amplitude A cause the wavefront to distort. However it is often, and in this case, assumed that astronomical observations will only occur under weak turbulence conditions, in which the effects of scintillation are negligible, (Roddier, 1981). For this reason A remains constant at 1, and we assume geometrical propagation (Fried, 1982). If the wavefront is incident on a layer of turbulence at height h and thickness δh, the wavefront undergoes a phase shift $\Delta\phi(x)$ due to the varying refractive indices $n(x,h)$:

$$\Delta\phi(x) = k\int_{h}^{h+dh} n(x,y,z)dz \tag{2.17}$$

where k is the wave number. For a turbulent phase screen of width w with a Von Karman PSD, the spatial power spectrum is:

$$\Phi_n(\kappa) = 2.54 \times 10^{-4} \delta^2 L_0^2 (1 + \frac{\kappa^2}{\kappa_0^2})^{-11/6} \tag{2.18}$$

where κ is a spatial frequency variable and $\kappa_0 = \frac{2\pi}{L_0}^2$ (Tyson, 2000). δ^2 is the variance of the refractivity fluctuations:

$$\delta^2 = \frac{C_n^2}{1.9 L_0^{2/3}}. \tag{2.19}$$

To recreate Kolmogorov turbulence, L_0 can be set to infinity. By multiplying the power spectrum by Gaussian white noise, a random spectral distribution is generated that retains the correct statistics. Gaussian white noise is a complex field with zero mean and unit variance, expressed as $(r(\kappa) + ir'(\kappa))$. From here, a phase screen $P(x, y)$ can be defined as:

$$P(x, y) = c\mathbb{R}\{\mathbb{F}[\sqrt{\Phi(\kappa)}((r(\kappa) + ir'(\kappa))]\} \tag{2.20}$$

$\mathbb{RF}$ here represents a real Fourier Transform (FT) and c is a scaling factor which means the strength of the turbulence in the phase screen can be defined in terms of r_0:

$$c = \frac{0.1517}{\sqrt{2}} (\frac{w}{r_0})^{5/6} \tag{2.21}$$

The phase screen can be expressed as a physical path difference (in metres), as opposed to radians of phase at a specific wavelength. This allows the same phase screen to be defined independently of wavelength.

Due to the Fast Fourier Transform (FFT) the phase screen will be periodic. This means the width W should be significantly larger than the aperture of the telescope. W should also be larger than L_0, or the turbulence from the low order modes will be undervalued. Figure 2.1 gives examples of phase screens generated using the python package AOTools (Townson et al., 2019). The two phase screen are generated using the same phase sampling and r_0, but have different outer scales.

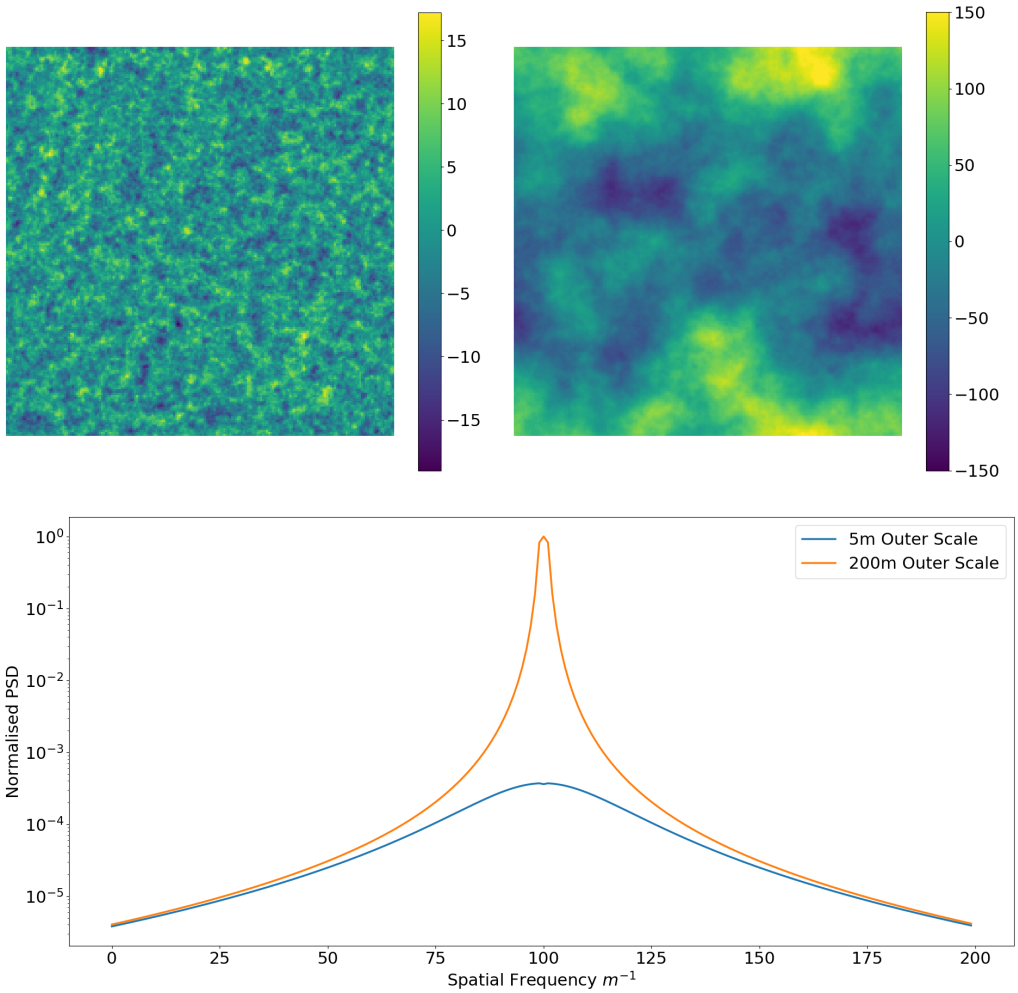


Figure 2.1: Example simulated phasescreens in radians: Von Karman turbulence with a 5m outer scale (left) and 200m outer scale (right). Both have the same r_0 and phase sampling, but the larger L_0 shows a wider variation in amplitude. Both phase screens represent 200x200m and spatial scale. Below: a normalised 1-D version of the PSDs used to generate both phase screens

The effect of a small outer scale is indicated in the small pockets of turbulence across the screen. For the larger outer scale the eddies stretch across the phase screen. In addition, note the changes in magnitude of the two phase screens, the smaller L_0 has suppressed the peak to valley optical path difference.

2.1.1.1 Infinite Phase Screens

The method detailed in 2.1.1 will generate a phase screen of finite size. In order to simulate frozen flow over the telescope, the aperture would instead have to move over the phase screen. This can be particularly problematic for modelling larger aperture telescopes due to the size of the phase screens that become necessary. As an example, an ELT-sized aperture of 39 m given 1024x1024 pixels, a one-minute exposure with a wind speed of 10 m/s would require a 20480x20480 screen. If using 64-bit data, this corresponds to over 3Gb of storage for a phase screen corresponding to one layer of turbulence for one minute.

As an alternative, "infinite" phase screens can be used (Assémat et al., 2006). Initially a finite phase screen is created in a similar fashion as discussed in section 2.1.1. The last two columns of the screen can be used as a stencil to create a new column of the phase screen at the same time as the first column is removed. This allows the array to stay the same size and the turbulence to appear to move across the phase screen; meaning the telescope aperture can also remain still. This kind of stencilling is sufficient for Von Karman turbulence, but for Kolmogorov turbulence the sampling of the original phase screen is more complex (Fried and Clark, 2008). This includes an additional reference point, near but not in the 2 columns used as the stencil, which can be used to ensure the Kolmogorov statistics from the structure function remain constant across the original phase screen and the new columns. The use of infinite von Karman or Kolmogorov phase screens allows for ELT scale simulations as well as very long exposure images to be generated.

Following the method presented in (Assémat et al., 2006), we define the stencil as

vector $\mathbf{Z}$, containing two columns of the initial phasescreen. This relates to the new phase data $\mathbf{X}$ as:

$$\mathbf{X} = \mathbf{A}\mathbf{Z} + \mathbf{B}\beta \tag{2.22}$$

where $\mathbf{A}$ and $\mathbf{B}$ are both matrices, and β is a random vector with Gaussian statistics. To find matrix $\mathbf{A}$ first multiply by $\mathbf{Z}^T$ and take the average over time. This reduces to:

$$\langle \mathbf{X}\mathbf{Z}^T \rangle = \mathbf{A}\langle \mathbf{Z}\mathbf{Z}^T \rangle \tag{2.23}$$

as β and $\mathbf{Z}$ are uncorrelated. From here we find $\mathbf{A}$:

$$\mathbf{A} = \langle \mathbf{X}\mathbf{Z}^T \rangle \langle \mathbf{Z}\mathbf{Z}^T \rangle^{-1} \tag{2.24}$$

The transpose of $\mathbf{A}$ is therefore:

$$\mathbf{A}^T = \langle \mathbf{Z}\mathbf{Z}^T \rangle^{-1} \langle \mathbf{Z}\mathbf{X}\mathbf{T}^T \rangle \tag{2.25}$$

as $\langle \mathbf{Z}\mathbf{Z}^T \rangle$ and its inverse is symmetric.

Next to find matrix $\mathbf{B}$, we use equation 2.22 and multiply by $\mathbf{X}^T$, giving:

$$\mathbf{X}\mathbf{X}^T = \mathbf{A}\mathbf{Z}\mathbf{Z}^T\mathbf{A}^T + \mathbf{A}\mathbf{Z}\beta^T\mathbf{B}^T + \mathbf{B}\beta\mathbf{Z}^T\mathbf{A}^T + \mathbf{B}\beta\beta^T\mathbf{B}^T \tag{2.26}$$

Taking the average of equation 2.26 and remembering β and $\mathbf{Z}$ are uncorrelated we get:

$$\langle \mathbf{X}\mathbf{X}^T \rangle = \mathbf{A}\langle \mathbf{Z}\mathbf{Z}^T \rangle \mathbf{A}^T + \mathbf{B}\mathbf{B}^T \tag{2.27}$$

$$\mathbf{B}\mathbf{B}^T = \langle \mathbf{X}\mathbf{X}^T \rangle - \mathbf{A}\langle \mathbf{Z}\mathbf{Z}^T \rangle \mathbf{A}^T \tag{2.28}$$

Substituting in equation 2.25 gives:

$$\begin{aligned} \mathbf{B}\mathbf{B}^T &= \langle \mathbf{X}\mathbf{X}^T \rangle - \mathbf{A}\langle \mathbf{Z}\mathbf{Z}^T \rangle \langle \mathbf{Z}\mathbf{Z}^T \rangle^{-1} \langle \mathbf{Z}\mathbf{X}^T \rangle \\ &= \langle \mathbf{X}\mathbf{X}^T \rangle - \mathbf{A}\langle \mathbf{Z}\mathbf{X}^T \rangle \end{aligned} \tag{2.29}$$

$\mathbf{B}\mathbf{B}^T$ is symmetric so can be re-expressed as:

$$\mathbf{B}\mathbf{B}^T = \mathbf{U}\mathbf{W}\mathbf{U}^T \tag{2.30}$$

using a singular value decomposition where $\mathbf{U}$ is the eigenvectors of $\mathbf{BB}^T$ and the diagonal values of $\mathbf{W}$ are its eigenvalues. As $\mathbf{W}$ is a diagonal matrix $\mathbf{W} = \mathbf{LL}^T$, where the diagonal elements of $\mathbf{L}$ are the square roots of the eigenvalues of $\mathbf{BB}^T$. Substituting this into equation 2.30 gives:

$$\begin{aligned} \mathbf{BB}^T &= \mathbf{ULL}^T\mathbf{U}^T \\ &= (\mathbf{UL})(\mathbf{UL}^T) \end{aligned} \tag{2.31}$$

and therefore:

$$\mathbf{B} = \mathbf{UL} \tag{2.32}$$

By substituting equations 2.24 and 2.32, representing matrices $\mathbf{A}$ and $\mathbf{B}$, back into equation 2.22, we can now calculate new phase data and extend the initial phase screen. This process can be repeated indefinitely, allowing for longer exposure times at smaller memory costs. Unless stated otherwise, the infinite phase screens from the AOTools package have been used.

2.1.2 Multi-Layered Profile

Phase screens describe a thin layer of atmosphere, so often multiple phase screens are used to model a full atmosphere. Each layer within a model atmosphere will be given a different altitude and strength J. J is the strength of the turbulence with height - $C_n^2 dh$ in units $m^{1/3}$. Measuring the turbulence as a function of height can be achieved by sending a balloon up through the atmosphere to measure the vertical temperature profile, or using ground based instruments (Sarazin et al., 2013).

Balloon sounding measurements of the temperature structure function are still used in present day, but they are impractical as a permanent measuring technique at astronomical sights as they cannot measure different altitudes concurrently (Bufton et al., 1972) (Azouit and Vernin, 2005). Weather balloons are of particular help when attempting to measure the atmospheric turbulence in inaccessible locations, such as the south pole (Azouit and Vernin, 2005).

Ground based techniques for measuring the atmospheric turbulence as a function of height include the SLOpe Detection And Ranging (SLODAR) method (Wilson, 2002). This can be used in real time and is ideal for astronomical sites. By using two stars, the overlap between the turbulence seen by each line of sight will decrease with increasing altitude, meaning $C_n^2(h)$ can be found from the correlation in the wavefront slopes(Butterley et al., 2006). SCIDAR uses a similar triangulation methodology but with the scintillation patterns for the two stars imaged onto a single detector Vernin and Roddier (1973). Stereo-SCIDAR builds on this by using one camera for each of the double stars, making it up to twelve times more sensitive than a traditional SCIDAR approach (Shepherd et al., 2014). These methods may utilise a Shack-Hartmann wavefront sensor (Platt and Shack, 2001). This uses a series of lenslets to split the incoming light into multiple different spots. The movement in each of these subimages demonstrates the changes in the phase local to each lenslet. When combined together this is indicative of the total wavefront seen by the telescope.

As measuring the atmospheric turbulence as a function of height requires real on-sky data, atmospheric profiles are often site specific. Here we use the median ESO profile, Marchetti (2015), generated from data from Paranal and Cerro Armazones. It defines 35 different layers from 30 m to 26500 m in altitude, and defines the Fried parameter as 0.157 m and the outer scale as 25 m at 500 nm. Table 2.1 shows the heights, speeds and relative strengths of this profile, but the directions of the individual layers .

We note that 50% of the strength of atmospheric turbulence occurs in the first 4 layers, with heights up to 200m. This is typical at astronomical observatories (Tokovinin et al., 2003).

Layer	Height (m)	Wind (m/s)	$\%J$
1	30	5.5	24.2
2	90	5.5	12
3	150	5.1	9.68
4	200	5.5	5.9
5	245	5.6	4.73
6	300	5.7	4.73
7	390	5.8	4.73
8	600	6	4.73
9	1130	6.5	3.99
10	1880	7	3.24
11	2630	7.5	1.62
12	3500	8.5	2.6
13	4500	9.5	1.56
14	5500	11.5	1.04
15	6500	17.5	1
16	7500	23	1.2
17	8500	26	0.4
18	9500	29	1.4
19	10500	32	1.3
20	11500	27	0.7
21	12500	22	1.6
22	13500	14.5	2.59
23	14500	9.5	1.9
24	15500	6.3	0.99
25	16500	5.5	0.62
26	17500	6	0.4
27	18500	6.5	0.25
28	19500	7	0.22
29	20500	7.5	0.19
30	21500	8	0.14
31	22500	8.5	0.11
32	23500	9	0.06
33	24500	9.5	0.09
34	25500	10	0.05
35	26500	10	0.04

Table 2.1: Table describing height above telescope, speeds and strengths as a percentage $\%J$ of the ELT 35 layer atmospheric profile. Taken from Marchetti (2015)

2.1.2.1 Measuring Wind Velocity

The velocity of the atmosphere as a function of height is not routinely measured at observatories. It can be physically measured, often by using weather balloons, but these take time to rise through the atmosphere, so cannot measure the velocity of all heights simultaneously or continuously - the NOAA balloons take 85 minutes to reach an altitude of 25km (Osborn et al., 2016). Balloons will also drift with the wind, so cannot measure exactly along one line of sight. They are also limited in that they can only be succesfully launched when the ground layer moves at 5 ms^{-1} or less (Storvold et al., 1998). Martin et al. (2011) used a plane with a spiral flight pattern, but is only usable in wind speeds of 10 ms^{-1} or less.

An alternative is to calculate the wind velocity profile from the ground. This can be done using the stereo SCIDAR two star optical triangulation technique. By measuring the light from two stars there will be overlap across the atmosphere between the two lines of sight. Aberrations from where there is overlap will be seen by both and given the angular separation of the two sensors the atmospheric errors can be assigned to a specific height. By assuming frozen flow, changes in the cross covariance of the optical phase aberration are attributed to the velocity of the turbulence. Wind speed and direction measurements using stereo SCIDAR have been shown to match that of weather balloon data and forecast data in Osborn et al. (2016). The SCIDAR data can measure the wind velocity for the full height of the atmospheric profile, is only able to assign velocities to teh stronger turbulent layers. Osborn and Sarazin (2018) shows that meteorological models can be used to predict turbulence patterns. In comparison to SCIDAR data, the forecast predictions have an RMSE in the wind direction of 29 degrees and cannot produce a velocity dispersion with height.

Wind velocity profiling data is available for the ground layer. For example the Nordic Optical Telescope (NOT) has a database from its Meteorological Station at La Palma, which includes ground layer speed and direction in around 5 minute

intervals. This records wind direction with north as 0 degrees, and east as 90. Lombardi et al. (2007) analysed this NOT data from 1998 to 2004 and showed that this particular site is more likely to have ground layer wind along the east-west axis, but there is still a heavy spread in all other directions. A bias to westerly wind is to be expected due to the Coriolis effect - the wind is more likely to run parallel to the earth's rotation than perpendicular to it.

2.2 Zernike Modes

It is often convenient to describe optical aberrations using a modal decomposition into an orthogonal basis set over a numerical grid. Zernike developed a commonly used set of orthogonal modes to describe aberration within circular apertures. The eponymous Zernike polynomials are a complete, orthogonal set of polynomials defined on a unit circle. Low-order Zernike polynomial modes correlate to classical optical aberrations e.g. defocus, coma and astigmatism, and can also be used to represent the effects of atmospheric turbulence in an optical system (Noll, 1976). It is worth noting that as they are continuous, they may not be orthogonal over a discrete series of points within a circle. The infinite series of polynomials Z_n^m can be sorted by Noll conventions using their radial degree (n) and azimuthal frequency (m). The polynomials are given in polar coordinates ($X = r\sin\theta, Y = r\cos\theta$), expressed as:

$$Z_n^m(r,\theta) \pm iZ_n^{-m}(r,\theta) = R_n^m(r)exp(\pm im\theta) \tag{2.33}$$

this leads to:

$$\begin{aligned} Z_n^m(r,\theta) &= R_n^m(r)\cos m\theta \text{ for m} \geq 0 \\ Z_n^{-m}(r,\theta) &= R_n^m(r)\sin m\theta \text{ for m} < 0 \end{aligned} \tag{2.34}$$

where $0 \leq r \leq 1$, so the modes are limited to within a unit circle. The radial function $R_n^m(r)$ follows as:

$$R_n^m(r) = \sum_{s=0}^{n-m/2} \frac{(-1)^s(n-s)!}{s![\frac{1}{2}(n+m)-s]![\frac{1}{2}(n-m)-s]!} \tag{2.35}$$

Zernike polynomials are commonly described using a single index as Z_j, where j increases by radial order and then azimuthal order ($j = [n(n + 2 + m]/2$). By ordering in this fashion, the polynomials split into symmetric and antisymmetric modes for even and odd values of j respectively:

$$Z_{evenj} = \sqrt{n+1}R_n^m(r)\sqrt{2}\cos m\theta \tag{2.36}$$

$$Z_{oddj} = \sqrt{n+1}R_n^m(r)\sqrt{2}\sin m\theta \tag{2.37}$$

$$Z_j = \sqrt{n+1}R_n^0(r) \text{ for m} = 0 \tag{2.38}$$

There are other ways of numbering the Zernike polynomials for example OSA or Fringe numbering. Fringe indexing is often used in optical design. As is common for work in astronomical adaptive optics, this work will follow the Noll conventions. The first 6 modes are shown in Table 2.2, listed with their radial and azimuthal orders and their classical aberration names.

Figure 2.2 shows the Zernike modes sorted into radial order and by same azimuthal order in an attempt to highlight the differences in the indexing. A series of the polynomials with different weights or coefficients a_j can be used to describe an aberrated wavefront $W(R\rho, \theta)$ passing through a optical system with radius R where $\rho = r/R$:

$$W(R\rho, \theta) = \sum_j a_j Z_j(\rho, \theta) \tag{2.39}$$

Z_j	n	m	Equation	Classical Aberration
Z_1	0	0	1	Piston
Z_2	1	-1	$2r\cos\theta$	Tip
Z_3	1	1	$2r\sin\theta$	Tilt
Z_4	2	0	$\sqrt{3}\ (2r^2$ -1)	Defocus
Z_5	2	-2	$\sqrt{6}\ r^2\ \sin 2\theta$	Astigmatism
Z_6	2	2	$\sqrt{6}\ r^2\ \cos 2\theta$	Astigmatism

Table 2.2: Table describing equations for the first six low order Zernike polynomials and their equivalent classical optical aberrations, generated using AOTools (Townson et al., 2019)

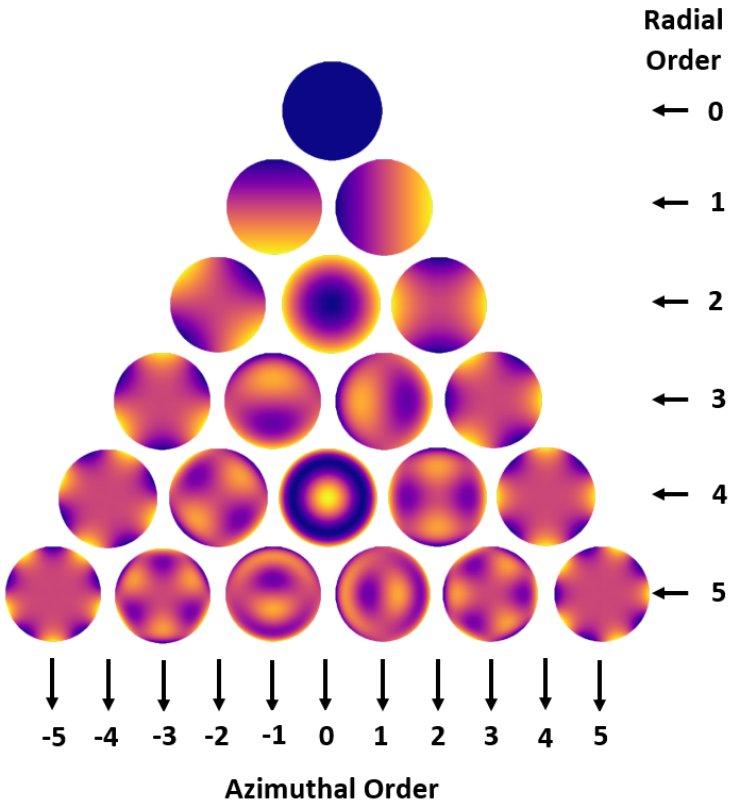


Figure 2.2: First 21 Zernike modes, sorted by radial order and azimuthal order.

Naturally, the larger the Zernike coefficient, the higher the contribution from that particular mode to the aberration on the system. The Zernike coefficients are independent of the number of polynomials used to describe a wavefront as the modes are orthogonal. For example, the coefficients for the first n modes will not change if more polynomials are used.

The Noll paper also derives the expected variances for the Zernike coefficients for Kolmogorov turbulence. These can be used as a verification for our simulation, running a test case of $D = r_0$. Figure 2.3 shows that simulation variances still follow in the same step like pattern and at very similar scales for larger outer scales, building confidence that the simulation is working as intended. As seen in Winker (1991), a finite outer scale will attenuate the Zernike variances, particularly the lower order modes. Note that even for $L_0/D = 10$, tip and tilt are still much lower than the Noll variances but the high order modes do match in scale.

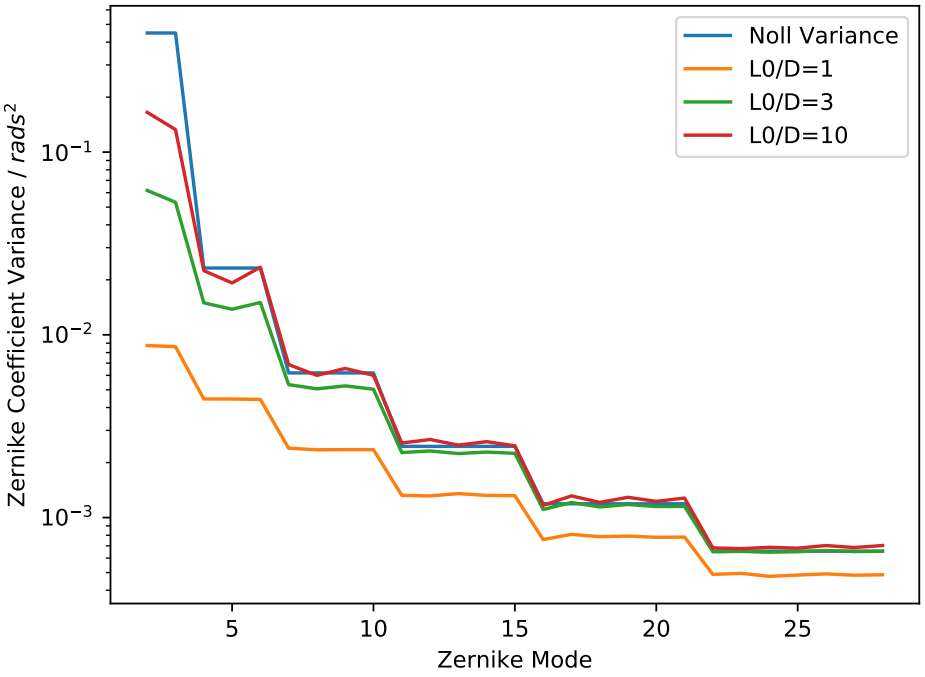


Figure 2.3: A comparison of Zernike coefficient variances from simulation for differing $L0/D$ values, excluding piston. A larger outer scale will converge in to the Noll variances, which assume Kolmogorov turbulence.

The results in figure 2.3 are produced using Zernike modes that have been orthonormalised using the Gram-Schmidt process (Swantner and Chow, 1994). Without this, using larger outer scales, modes from higher radial orders begin to exhibit odd behaviours. As the outer scale gets larger, spikes begin to appear in the instantaneous Zernike variance for positive and even azimuthal order modes, reaching an order of magnitude higher than other modes in that radial order. To reduce this the phase used to generate the covariances should be zero mean. To achieve this we can utilise the Gram-Schmidt process. Afterwards, the variances appear as expected, as seen in figure 2.4. All Zernike coefficients in this work have been generated with Gram-Schmidt Zernike modes.

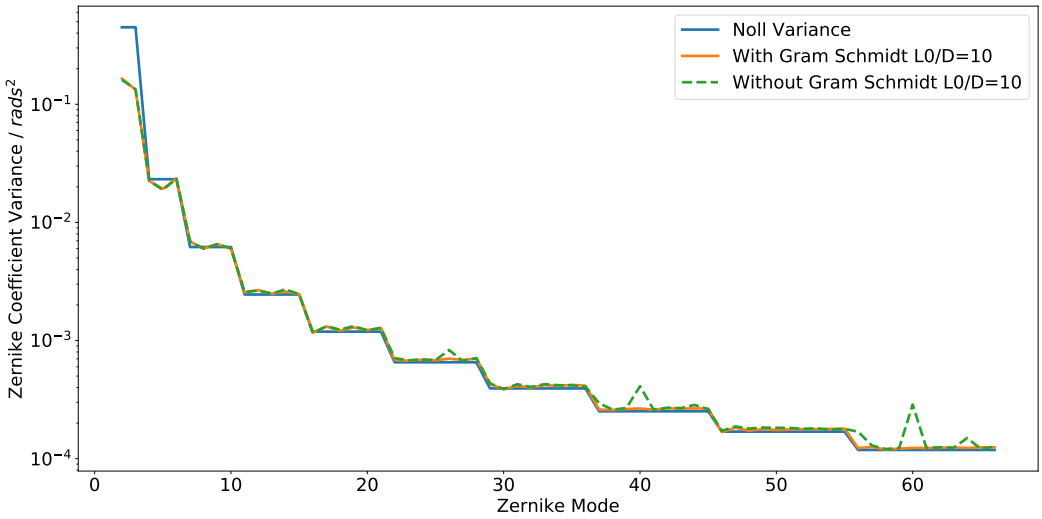


Figure 2.4: Instantaneous Zernike Variance for $L0/D = 10$ compared to Noll variance excluding piston. Without the use of Gram-Schmidt orthonormalisation, spikes occur for modes with azimuthal order +4 and +8.

2.2.1 Alternative Modes

As most optical set ups use circular apertures, they are a useful tool when describing an aberrated wavefront. Zernikes are not the only set of polynomials that can be used to describe optical aberrations. Systems with alternative aperture shapes may use different series, but Zernikes still hold the advantage of corresponding to classical optical aberrations (Lakshminarayanan and Fleck, 2011). Systems that use annular apertures may choose Karhunen-Loeve modes as an alternative. A third fewer KL modes than Zernike modes are required to produce an equivalent match to a wavefront Cannon (1996).

Regular Zernikes are not orthogonal over non circular apertures, but basis changes, such as the Gram-Schmidt method, can be performed to alter the series for annular or elliptical systems (Swantner and Chow, 1994).

2.3 Long exposure simulation method

In order to explore the effects of long term phase averaging a Monte Carlo simulation of atmospheric turbulence was used to model atmospheric wavefront errors as a function of wavefront sensor averaging time. The model uses infinite phase screens (Assémat et al., 2006), to allow for long exposures to be simulated without the prohibitive computer memory requirements.

For an individual aperture observing a single layer of turbulence, a section of the phase screen at time t which is the same size as the aperture p, q is selected; $\phi^{'}_{screen}(t, p, q)$, where p and q are pixel coordinates. The instantaneous Zernike coefficient $a^{'}_{j}(t)$ for an individual mode is calculated by multiplying the Zernike mode $Z_j(p, q)$ by the phase screen and summing across the aperture

$$a'_j(t) = \frac{\sum_p \sum_q \phi'_{screen}(t, p, q) Z_j(p, q)}{A}, \tag{2.40}$$

where A is the area of the circular pupil. Zernike modes have been indexed and normalised using the Noll convention (Noll, 1976).

The phase screen is translated across the telescope aperture in a series of small steps, and the resulting instantaneous coefficient is calculated repeatedly at different values of time t. The long exposure Zernike coefficient for an averaging time of τ can then be calculated as

$$a_j(\tau) = \frac{1}{N_t} \sum_{t=0}^{N_t(\tau)} a'_j(t), \tag{2.41}$$

where N_t is the number of time steps in the period τ. The mean Zernike coefficient represents the magnitude of phase aberration represented by a mode within an exposure time τ. To account for the finite exposure time of the simulation, this mean coefficient can be calculated for different realisations of the atmosphere by generating new phase screens with the same input parameters. The variation of

τ is how this thesis tests the effects of an increasing but finite averaging time. As the Zernike coefficients are zero-mean the RMS of the Zernike coefficient σ_j is calculated as

$$\sigma_j(\tau) = \sqrt{\langle |a_j(\tau)|^2 \rangle}. \tag{2.42}$$

The RMS $\sigma(\tau)$ represents the amount of phase aberrations sampled by the mode, integrated over a finite exposure time. For a simulation with many layers, the Zernike variances from the individual layers are summed together to create a total variance and then the RMS is calculated. Geometrical propogation is assumed (Fried, 1982). The results are also weighted for the strength of each layer $\%J$, given its height as discussed in section 2.1.1, giving a total $\sigma_{tot}(\tau)$ of:

$$\sigma_{tot} = \sqrt{\sum_{h=0}^{h=max} (V_j \cdot \frac{\%J}{100})} \tag{2.43}$$

Chapter 3

Effects of wind velocity profiles on turbulence-induced quasi-static aberrations

3.1 Introduction

In this chapter, we investigate the impact of wind velocity on the characteristic timescales required for turbulence averaging, focusing on next generation ELT scale telescopes. An AcO system aims to correct for errors inherent to the telescope. Ideally the reference centroids should correspond to a flat wavefront. This can be achieved on-sky by averaging over a sufficiently long enough period of time so that the atmospheric effects will average to zero, leaving only the instrument error for the AcO system to correct. The time this takes scales strongly with telescope diameter, which may cause problems for ELT-scale apertures. Part of this work was previously presented in Gill et al. (2022).

A simulation was used to explore the decay rates of time-averaged atmospheric turbulence residuals under frozen flow conditions. This chapter first shows that the RMS of Zernike coefficients averaged across each radial order decreases with exposure time with an approximate -0.5 power law, in line with the published

results from Gordon et al. (2011). However, our model shows that there is a significant deviation in the power law exponent when individual Zernike modes within the same azimuthal pair are considered. For example, for a 39 m telescope, Fried parameter of 0.14 m and outer scale of 39 m and a wind speed of 10 m/s, the times taken for Zernike modes 5 and 6 to reach a residual error of $\lambda/20$ are 10 and 64 seconds in the K-band and 40 and 1016 seconds in the V-band respectively.

This chapter shows via simulation that the characteristic timescale needed for a time-averaged Zernike mode to achieve a specified wavefront variance is dependent upon wind direction, meaning that both the C_n^2 and wind velocity profiles can have a large effect on the modal variance of observed quasi-static aberrations. We initially explore the effects of varying wind direction on individual modes using a single layer turbulence model (section 3.2.2). This was then extended in section 3.3 to incorporate some angular dispersion of wind direction in a multi-layer atmosphere. We show that a larger angular dispersion is equivalent to averaging over the azimuthal order. This means that for atmospheres where there is a wide dispersion of wind directions the characteristic averaging timescale tends back to the results presented in Gordon et al. (2011); whereas for a bulk wind direction the timescale is closer to the results of a single layer.

As the turbulence closer to the ground is typically stronger, varying the directions of the lowest layers in our model has a clear effect on results generated with the same angular separation between layers. A mean profile is generated looking at on sky data from a SCIDAR instrument, as used in Osborn et al. (2016). The multi-layered profile allowed us to investigate the effects of varying the wind speed of the ground layer in section 3.4. For example combined with ground wind speed statistics from the ESO Paranal ambient conditions database ESO (1998), our model predicts that for an ELT scale telescope 10% of the time mode 5 will take over an hour to reach a $\lambda/20$ at 500 nm threshold. Unless stated otherwise, all results presented in this chapter are for a 39 m diameter telescope with an atmospheric profile of 39 m outer scale and 0.157 m Fried parameter.

3.2 Single Layer Model

3.2.1 Azimuthally Averaged Zernike Modes

Following the method presented in section 2.3, we have explored the behaviour of the RMS WFE of the time averaged Zernike coefficients. Gordon et al. (2011) notes two regimes, separated by a break point $\tau_{bp} \approx D/v$ where D is the pupil diameter and v is the wind speed. This is recreated in figure 3.1. Initially the RMS WFE of the time averaged Zernike coefficients is near constant for short exposures and then decreases approximately as a power law after the break point.

Figure 3.1 presents $\sigma_n(\tau)$ for azimuthally averaged Zernike coefficients - the average of an azimuthal pair or a full radial order produces the same results. For long exposures, the WFS will sample different values of the Zernike modes as the atmosphere moves past. In a "heuristic analysis", Gordon assumes before and after τ_{bp} as uncorrelated, so the latter can be treated as a random walk process. The azimuthally averaged $\sigma_n(\tau)$ decreases approximately as a -0.5 power law - $\sigma_n(0)(\tau/\tau_{bp})^{-0.5}$ from Gordon et al. (2011). This is due to the rotational symmetry of the azimuthal orders averaging as the phase translates across. The exposure time needed to reach τ_{bp} will be different for different radial orders, shown in figure 3.1. The curves for the different radial orders are parallel but τ_{bp} is longer than for smaller modes.

The results in figure 3.1 are not affected by changes in wind direction, but are affected by wind speed. We demonstrate this dependency in figure 3.2, which shows the time averaged Zernike RMS WFE for radial order 2. The only change between the two lines is the speed at which the phase moves past the aperture - 5.1 ms^{-1} and 32 ms^{-1} respectively. These speeds have been chosen as they are the extremes shown in table 2.1 and are representative of possible wind speeds at the Paranal site. As discussed in Gordon et al. (2011) both scenarios start with the same Zernike coefficient RMS WFE and decrease with the same power law, as the same atmospheric statistics and pupil diameter have been used - but the

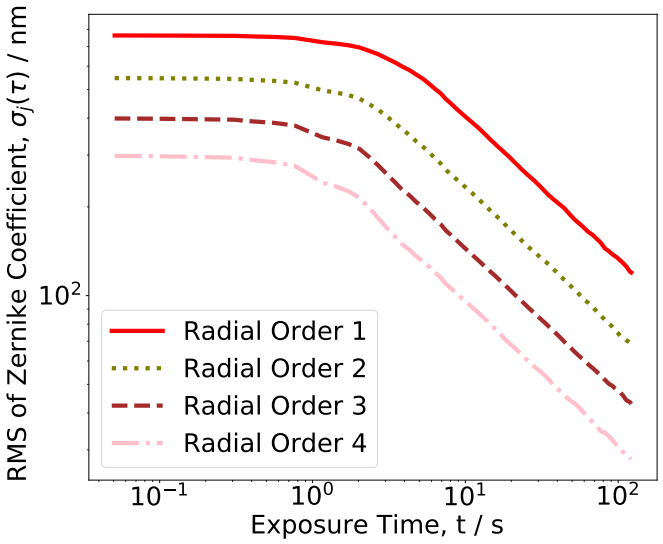


Figure 3.1: Time-averaged RMS WFE for frozen flow atmospheric turbulence for Zernike radial orders 1 through 4, showing the same gradient for each azimuthally averaged pair. This is a replication of the results from Gordon et al. (2011) using a single layer of von Karman turbulence with outer scale L_0 = 39 m and r_0 of 0.157 m moving at 10 ms^{-1} for an ELT sized aperture of diameter 39 m.

faster velocity breaks first. This can have large effects on averaging times if we wish to reach a specific threshold. The averaging times are linearly dependent on wind speed. For example, to reach a threshold of $\lambda/10$ will take six times longer for the slower wind speed for this particular case. To expand on these results we first continued with a single layer model but looked at individual Zernike modes.

3.2.2 Behaviour of Individual Modes

We started with a simple atmospheric model using a single phase screen of Von Karman turbulence with a Fried parameter r_0 of 0.157 m and outer scale L_0 of 39 m. Using the infinite phase screen function of aotools as discussed in section 2.1.1, we translated this past a 39 m aperture for 100 seconds with a wind speed $v = 10m/s$. For each step, we measured the instantaneous Zernike coefficients for the first 10 radial orders containing 66 modes in Noll convention. After repeating this process for different phase screens with the same statistics, we can calculated

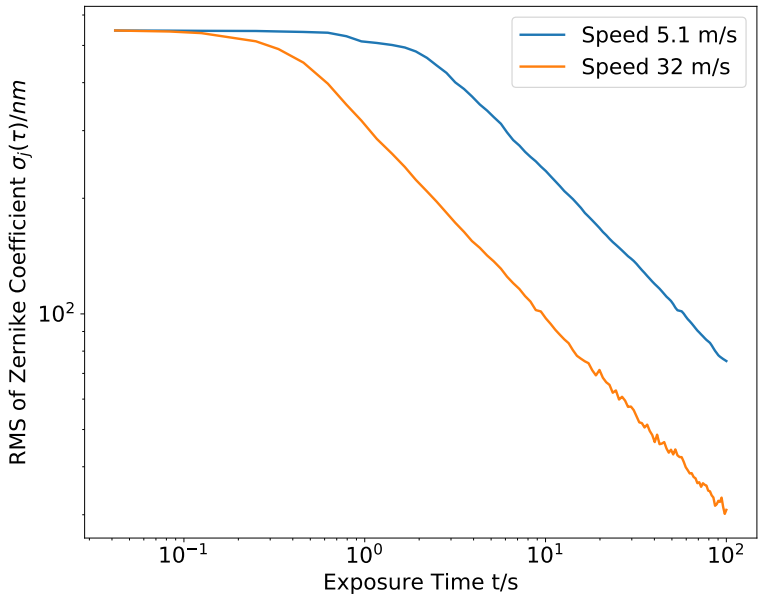


Figure 3.2: Dependence of the time-averaged residual wavefront error of mode focus versus averaging time for an ELT sized aperture using a single layer of von Karman turbulence with outer scale $L_0 = 39$ m and r_0 of 0.157 m with wind speeds of 5.1 ms^{-1} and 32 ms^{-1}, showing the break point's dependency on wind speed.

the Zernike modal distribution of time-averaged residual WFE as a function of averaging (or WFS exposure) time as shown in equation 2.42.

Figure 3.3 shows that the time taken for the averaging error of individual modes to reach a limit is also dependent on their azimuthal order. In the same way as the azimuthally averaged results shown in section 3.2.1, the averaging error for individual modes also follows the trend of constant RMS WFE for short exposures until reaching a break point at which they follow an approximate power law. The RMS WFE for Zernike coefficients within the same radial order are the same until the break point τ_{bp}, but have varying behaviour afterwards. In this logarithmic space a power law presents as a straight line. As well as the azimuthally averaged results from section 3.2.1, figure 3.3 shows that individual modes also act approximately as power laws. We have fit power laws to the Zernike RMS WFE after the break point. This is shown as the the fainter straight lines in the figure. We see that this approximate straight line fit is only valid for exposure times after the break point, so only these times are used to generate the fit. We fit gradients of -0.46 and -1 for

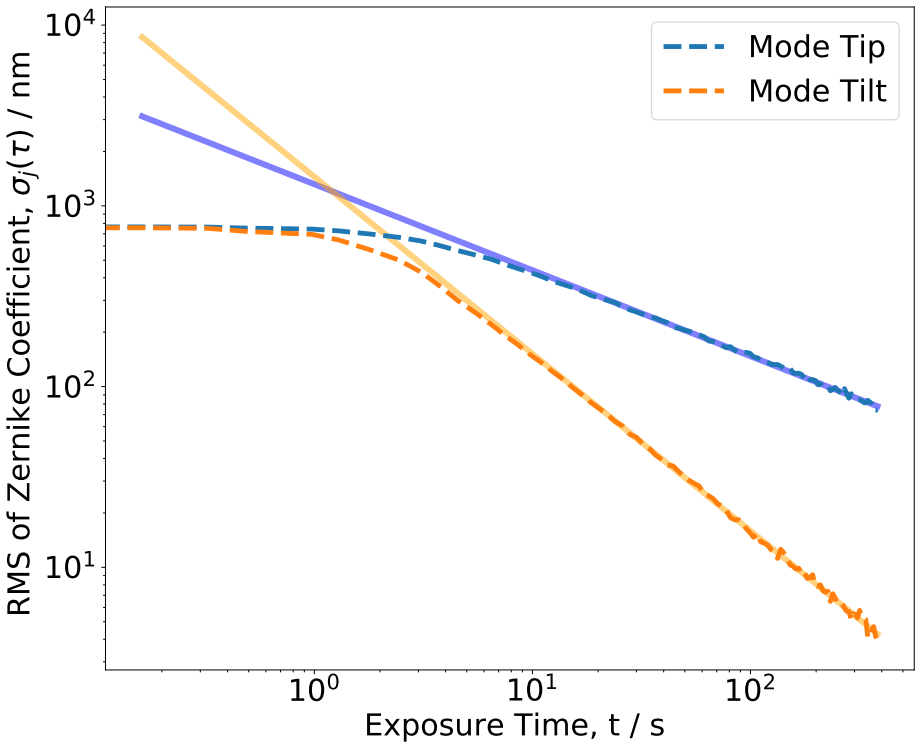


Figure 3.3: WFE for time averaged Zernike coefficients in nm RMS shown as a function of exposure time for Zernike Modes 2 and 3, tip and tilt, using a single layer of von Karman turbulence with outer scale $L_0 = 39$ m and r_0 of 0.157 m moving at 32 ms^{-1} for an ELT sized aperture of diameter 39 m.

modes tip and tilt respectively.

3.2.3 Effects of Varying Wind Direction

To explore the effects of wind direction on the results shown above in section 3.2.2, we repeated this process whilst applying a rotation matrix to the individual Zernike modes prior to calculating the instantaneous coefficients. In figure 3.4 we plot the same time-averaged residual WFE for the modes of radial order 2, with azimuthal orders 0 and ± 1, the only difference between each of the simulations being the wind direction.

The right and lower panels of figure 3.4 show that the approximate power law

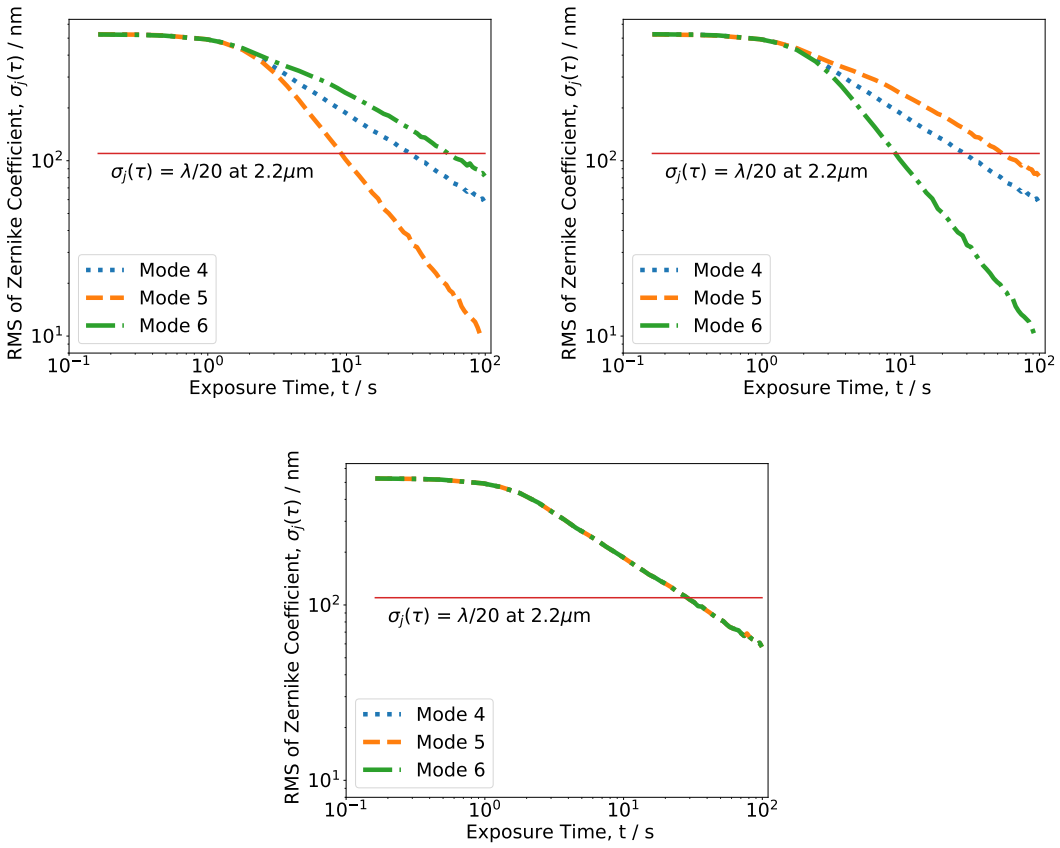


Figure 3.4: RMS WFE for time averaged Zernike coefficients in nm shown as a function of exposure time for Zernike Modes 4,5 and 6, (focus and astigmatisms) using a single layer of von Karman turbulence with outer scale $L_0 = 39$ m and r_0 of 0.157 m moving at 10 ms^{-1} for an ELT sized aperture of diameter 39 m. Panels left, right, bottom for wind directions 0, 45, 22 degrees respectively. Also shown on each panel is the threshold of $\lambda/20$ at 2200 μm (K Band).

for an individual mode is dependent on both its azimuthal order and the wind direction. This azimuthal order dependence stems from the rotational symmetry of the modes. Modes 4 and 5 have azimuthal order ± 2 respectively and are shown in figure 3.5 to demonstrate their rotational symmetry. It follows logically that the results produced by rotating mode 5 by 45 degrees should match that of mode 4 with no rotation and visa versa.

In addition after rotating both by 22 degrees, halfway between 0 and 45 degrees, the modes are symmetrical about the wind direction and so have identical RMS

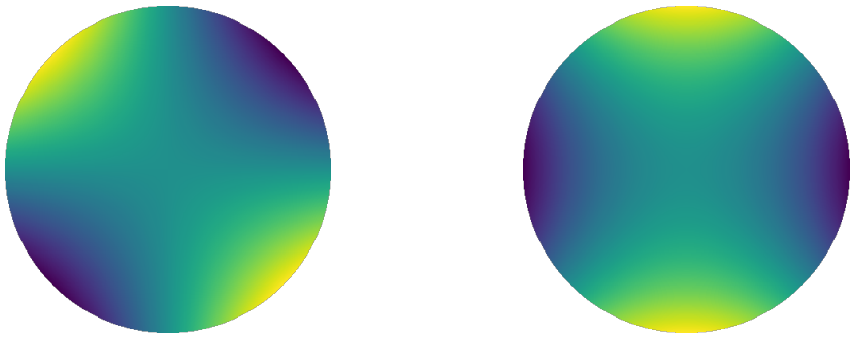


Figure 3.5: Zernike Modes 5 and 6, astigmatism 1 and 2 with no rotation applied.

WFE power laws. All three panels of figure 3.4 show that the averaging error for mode 4 (focus) is unaffected by wind direction. This is to be expected as mode 4 has zero azimuthal order and is always symmetric about the wind direction. Mode 4 reduces at an approximate -0.5 power law - the same as the azimuthally averaged results presented above. At 22 degrees when modes 3 and 5 align, they match the zero order time-averaged wavefront error power law. Using the same power law fit discussed in section 3.2.2, we plot the gradients of each mode for a full rotation of 0 through 90 degrees in figure 3.6. This clearly shows that the gradients of the three modes align at 22 degrees and that the astigmatisms switch positions over 45 degrees. The angular separation between the extremes in gradient can be expressed as $\frac{180}{|m|}$ degrees.

Extending this to higher order modes, at 0 degrees wind direction where the wind direction is aligned with the origin of the azimuthal angle of the Zernike mode, the slope gradients after the break point for all pairs of modes show the greatest difference. However the angle at which both modes are symmetric about the wind direction varies with azimuthal order. For example, the gradient of the slope after the break point for Zernike modes with azimuthal order 10 will switch places every 9 degrees, whereas tip and tilt would require a wind direction change of 90 degrees. The behaivour of azimuthal order 10 is shown in figure 3.7, and tip and tilt are plotted in 3.7. As a reference, the Zernike modes are shown in figures 3.9 and 3.10.

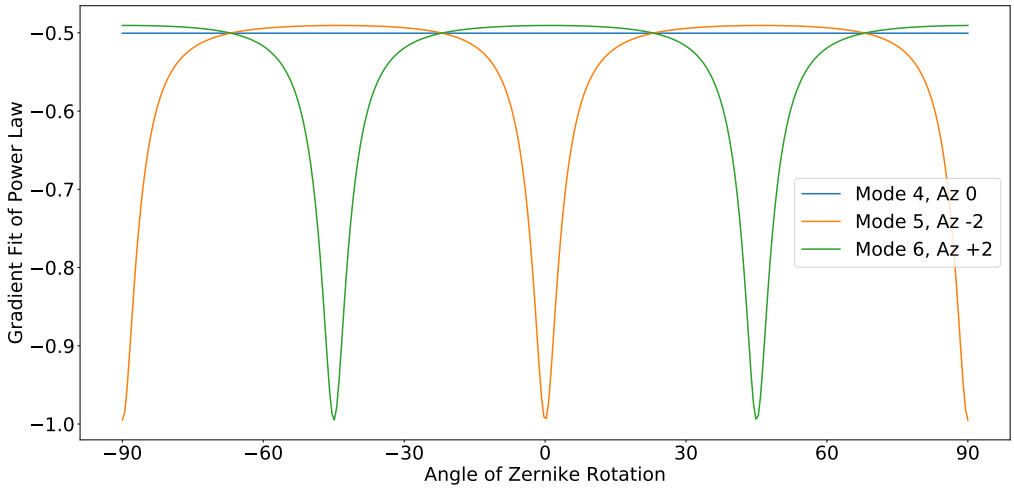


Figure 3.6: Power law gradient for radial order 2 as wind direction rotates from -90 through 90 degrees.

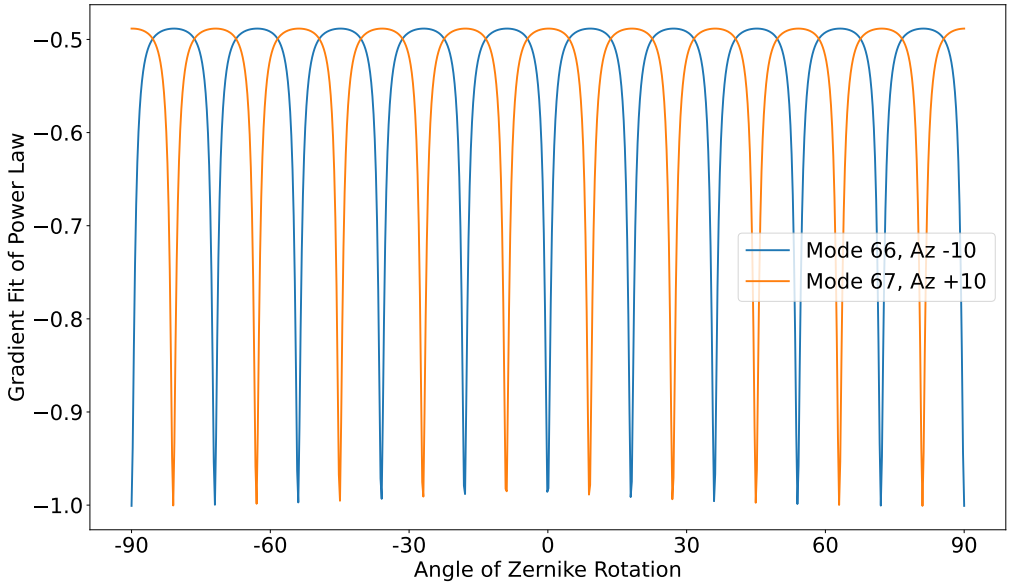


Figure 3.7: Power law gradient for an azimuthal pair ±10 as wind direction rotates from -90 through 90 degrees.

For a single phase screen this behaviour can be clearly explained, but for a more realistic multi-layered atmospheric model this would require the whole atmosphere to exhibit bulk motion in a single direction. In the next section we instead consider the case of a more complex atmosphere with multiple turbulent layers, each of which

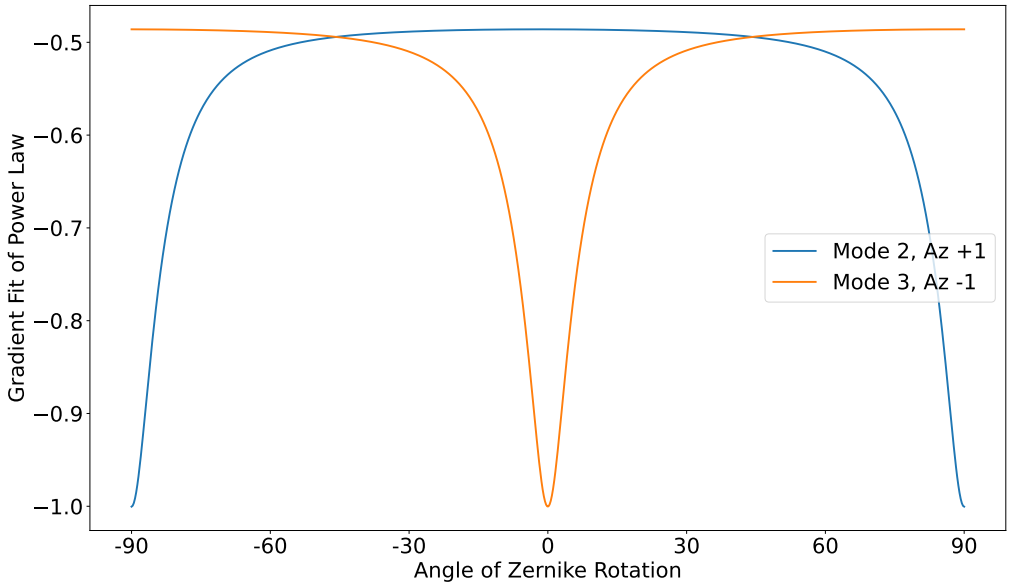


Figure 3.8: Power law gradient for an azimuthal pair ±1 as wind direction rotates from -90 through 90 degrees.

can have its own velocities.

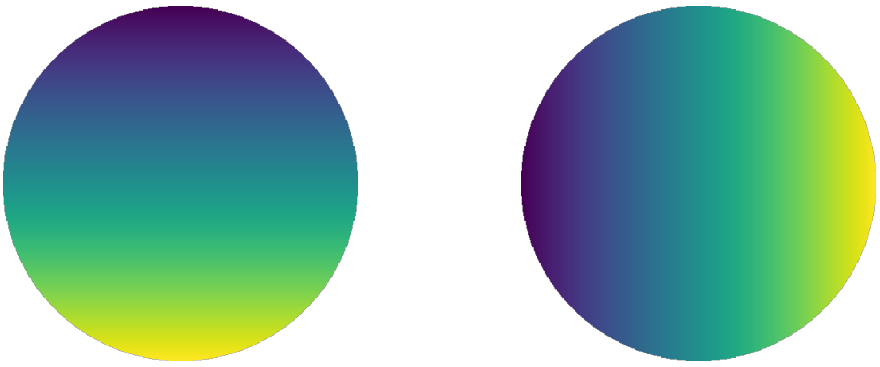


Figure 3.9: Zernike Modes 2 and 3, tip and tilt, with no rotation applied

3.3 Extending to Multiple Layers

To investigate the impact of wind velocity on the averaging time for more realistic atmospheres, we implemented 3 regimes:

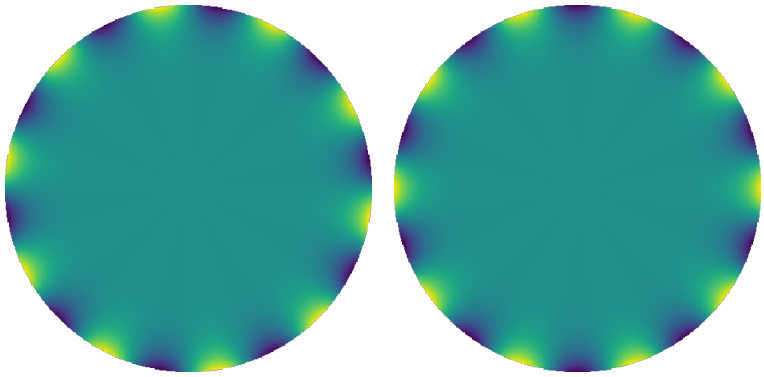


Figure 3.10: Zernike Modes 65 and 66, azimuthal order +-10 with no rotation applied.

1. The ESO 35 layer profile with vertical wind speed profile as described in section 2.1.2. Leaving the lowest layer translating at "0" degrees, a second direction was chosen for the highest layer. The wind direction of intermediate layers was evenly spread as a function of layer number. This gives helically varying wind directions between the lowest and highest layers. To find the wind direction θ_i for a layer i in our 35 layer profile:

$$\theta_i = \frac{1}{35-1}(i-1)\theta_{end} \tag{3.1}$$

   where θ_{end} is the chosen direction for the highest layer. This allowed for large angular dispersions and also limited ones that mimicked a bulk motion of the atmosphere.

2. We repeated the helical dispersion and also varied the direction of the lowest layer. This followed in much the same way as equation 3.1 but we also define a starting angle θ_{start} for the lowest layer,

$$\theta_i = \frac{1}{35-1}(i-1)(\theta_{end} - \theta_{start}) + \theta_{start} \tag{3.2}$$

3. Used SCIDAR data to produce a random wind velocity specified by a Gaussian distribution.

3.3.1 Helically Varying Wind Directions with Fixed Ground Layer

In this section we explore how wind directions of a multi-layered atmospheric profile affects the time averaged variance of Zernike modes. We kept the direction of the lowest layer fixed at 0 degrees - equivalent to the results shown in 3.2.2 - and helically vary the velocities of the 34 higher layers. Whilst previous results showed the evolution of Zernike variance with increasing exposure time for individual modes, here we show the variances for multiple modes after a set exposure time.

Figure 3.11 shows the time-averaged wavefront variance for individual Zernike modes after an exposure time of 100 seconds for the full 35 layer profile with all layers translating in one direction (at 0 degrees). The time-averaged wavefront variances for pairs of azimuthally-averaged Zernike modes are shown as the green lines for comparison. This result is akin to what we see in the first part of figure 3.4, all the modes within their azimuthal pairs are at their maximum offset from the azimuthal average.

A next step is to suggest that the atmosphere would have a bulk direction but with a small angular dispersion of wind velocities. To represent this we helically vary the layers through 45 degrees, as shown in figure 3.12. Also included in figure 3.12 is a "wind rose" to demonstrate the individual velocities of the full atmospheric profile. This shows that atmospheres with small angular dispersions can still give rise to very large differences in atmospheric turbulence residuals. All non-zero azimuthal-order modes have moved to be closer to the azimuthal average. Higher order being more effected by this small increase given their smaller angular symmetry. The comparison between figures 3.11 and 3.12 also shows that modes with zero azimuthal order and azimuthally averaged results are unaffected by wind direction.

Figure 3.13 shows the variances after 100 seconds given a helix between 0 and 360 degrees. We now see that all modes are very close to the azimuthal average as we are now equivalently averaging over wind direction. They are not identical results

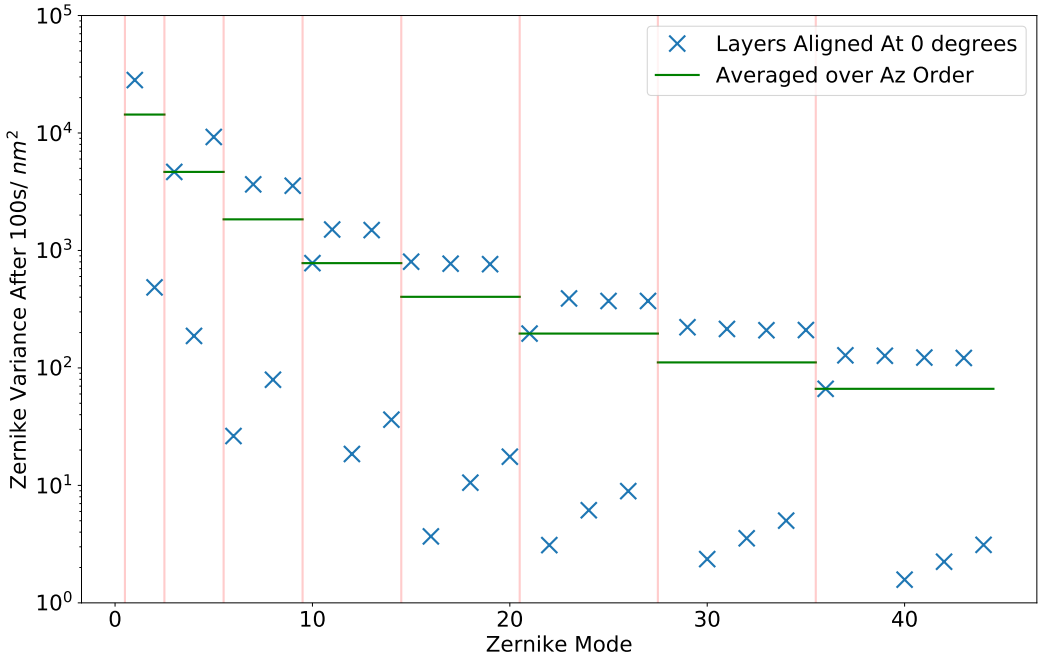


Figure 3.11: Modal breakdown of residual wavefront variance after integration time of 100 seconds for all layers aligned at 0 degrees. The red bars indicate different radial orders. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m, wind speeds detailed in table 2.1, for an ELT sized aperture of diameter 39 m.

though, given the atmospheric profiles discreet layers with individual speeds and weighting.

Individual layers have their own specific directions and strengths. Layers closer to the ground are more dominant so as the angular dispersion increases the variance does not tend smoothly to the azimuthally averaged case. This is seen in figure 3.14 which shows the time taken to reach a threshold of $\lambda/20$ at 550 nm for radial order two as the angular dispersion increased. The RMS WFE for the azimuthal pair never exactly average out to the zero order case, but their split does decrease as the angular dispersion increases. For a given wind speed profile, the wavefront averaging time can vary by an order of magnitude depending on the distribution of layer directions.

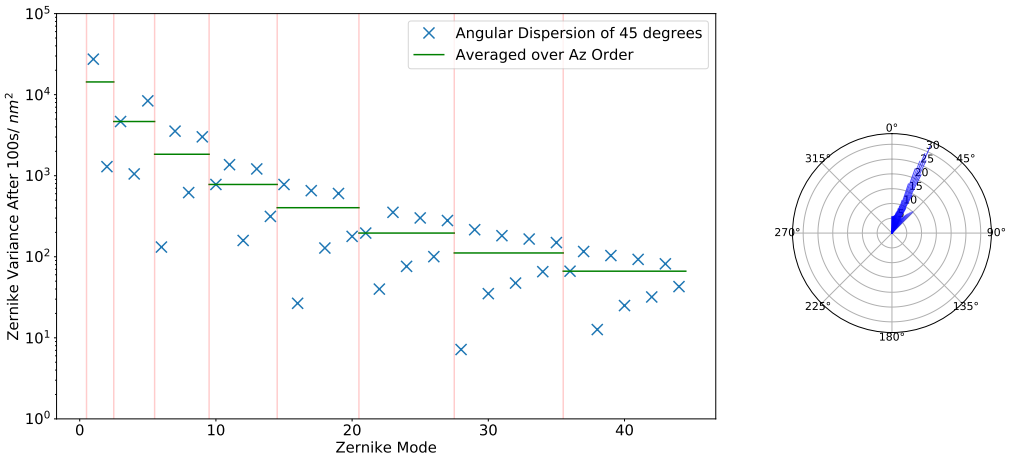


Figure 3.12: Modal breakdown of residual wavefront variance after integration time of 100 seconds for layers helically varied from 0 through 45 degrees. The azimuthally averaged variances are shown as the green lines. The polar plot shows the layers velocity distributions in ms^{-1}, starting with a surface layer aligned at 0 degrees, with respect to tip. The length of each line represents that layer's speed. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

3.3.2 Offsetting the Ground Layer

The results in the above section explore the effects of varying angular dispersion, but the helically varying profile starts at the ground moving at 0 degrees. The lowest three layers, below 200 m, hold just over half the total weight of the entire profile, so naturally the directions of these layers heavily affect the variance. In addition to varying the total angular dispersion of the whole wind helix we have also tested the effects of changing its starting angle. For example, starting with the lowest layer at 0 degrees and helically varying to the highest at 360 degrees has the same total angular dispersion as starting at 10 degrees and ending at 370 degrees; but will produce a different residual wavefront error.

By offsetting the wind velocity by just 10 degrees the distribution of modes appears less regular and closer to the azimuthally averaged case for an equivalent angular dispersion. Starting again with all layers aligned in a singular direction, figure 3.15 demonstrates how differently the individual modes are affected by wind direction. Here whilst only some modes are noticeably closer to the azimuthal average, the

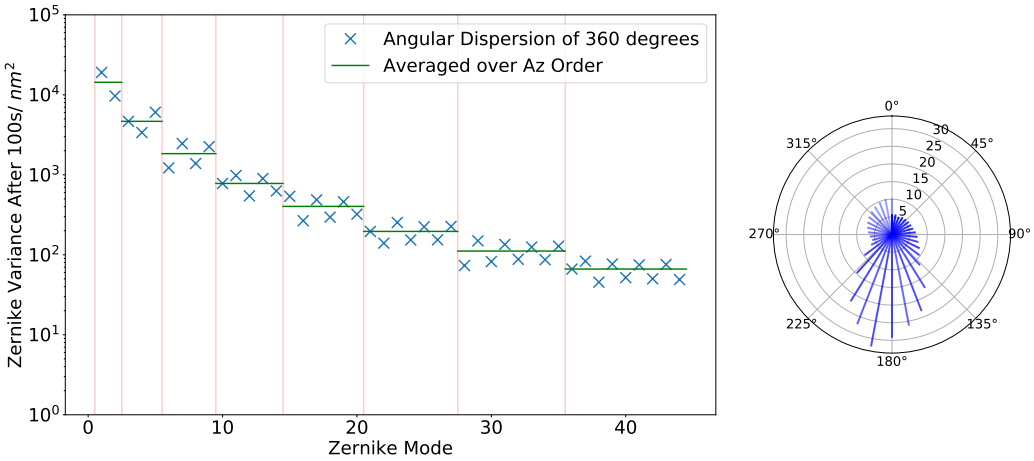


Figure 3.13: Modal breakdown of residual wavefront variance after integration time of 100 seconds for angular dispersion of all layers through 360 degrees. The azimuthally averaged variances are shown as the green lines. The polar plot shows the layers velocity distributions in ms^{-1}, starting with a surface layer aligned at 0 degrees, with respect to tip. The length of each line represents that layer's speed. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

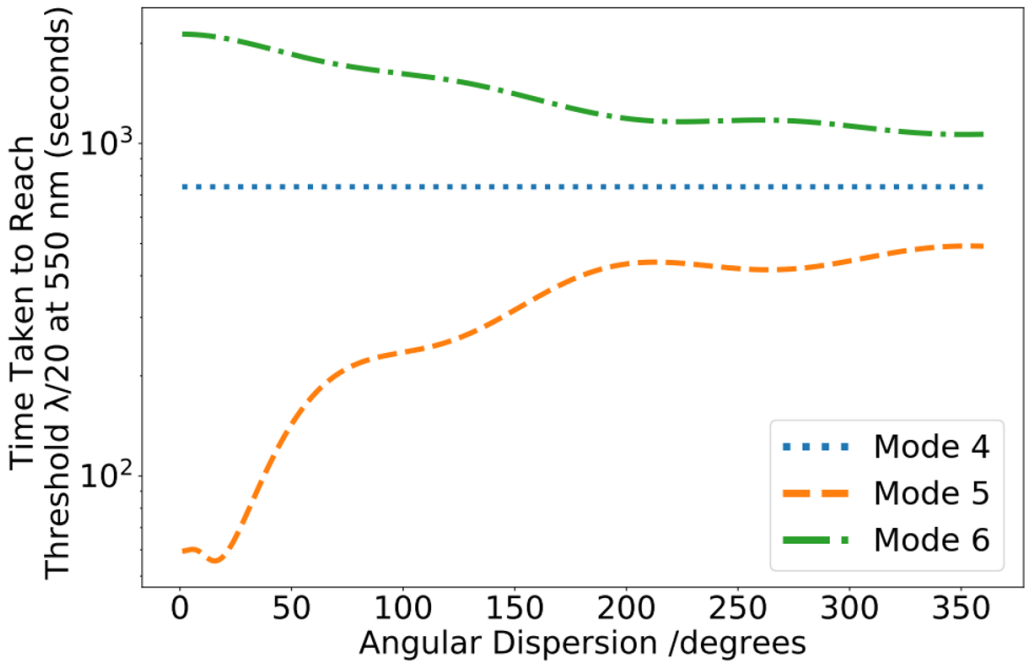


Figure 3.14: Time taken to reach an RMS WFE threshold of $\lambda/20$ at 550 nm for increasing angular dispersion, starting at 0 degrees. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

offset ground layers have a much clearer effect when considering angular dispersions.

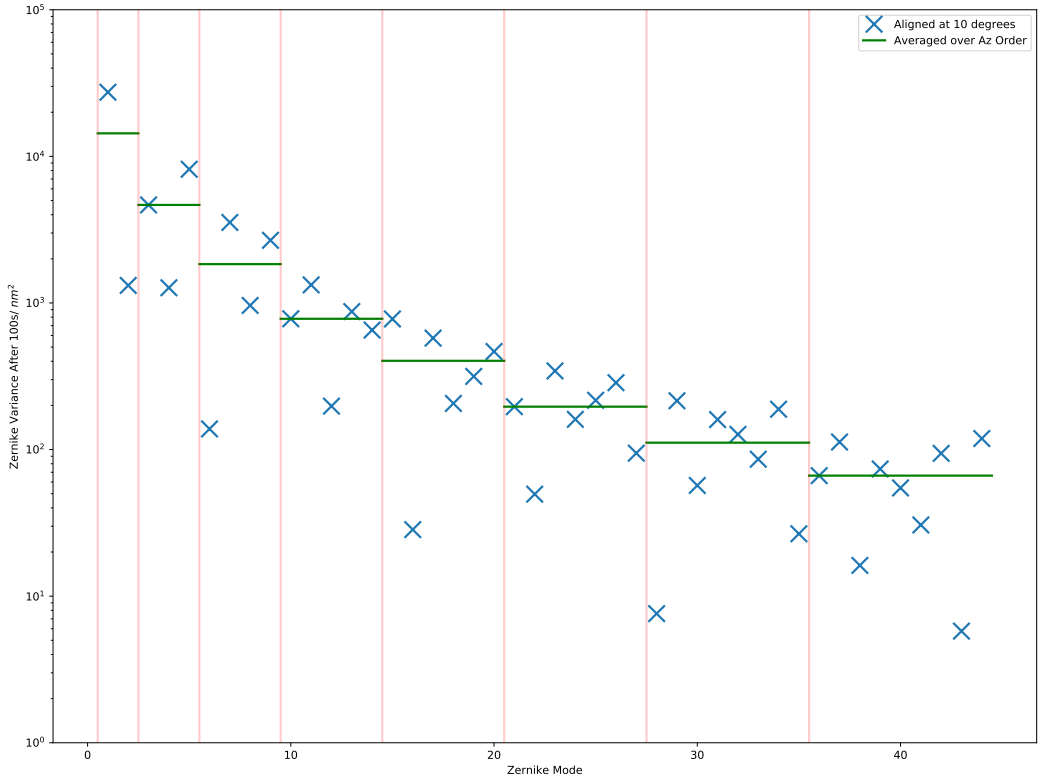


Figure 3.15: Modal breakdown of residual wavefront variance after integration time of 100 seconds for all layers aligned at 10 degrees. The azimuthally averaged variances are shown as the green lines. The red bars indicate different radial orders. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

Figure 3.16 shows a small dispersion of 45 degrees, but with the lowest layer moving at 10 degrees rotation. The comparison between figures 3.12 and 3.16 show the strong effect that the ground layer has on the overall results. The same small angular dispersion has a much smaller residual wavefront error when the ground layers change direction. Accurate profiling of wind direction at low altitudes would appear to be useful in reducing unnecessary averaging time - possibly by rotating the Zernike modes to a more optimal position. Consider the difference between figures 3.14 and 3.17: the x axis on both shows the same degree of angular dispersion, but the small additional offset of 10 degrees at the ground consistently affects the non-zero azimuthal order modes.

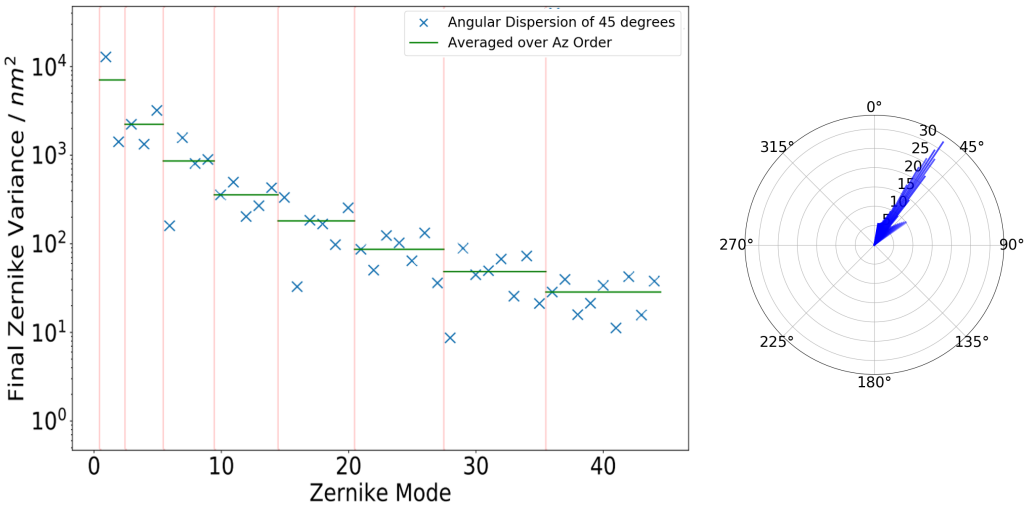


Figure 3.16: Modal breakdown of residual wavefront variance after integration time of 100 seconds for layers helically varied from 10 through 55 degrees. The azimuthally averaged variances are shown as the green lines. The polar plot shows the layers velocity distributions in ms^{-1}, starting with a surface layer aligned at 10 degrees, with respect to tip. The length of each line represents that layer's speed. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

Figure 3.17 shows the time taken to reach a threshold of $\lambda/20$ at 550 nm as the angular dispersion increases with the bottom layer at 10 degrees. This shows in quite a pronounced way that a larger spread of wind velocities does not always guarantee a smaller residual wavefront error. Here we see that radial order 2 is much closer to the azimuthally averaged case with an angular dispersion between 10 an 180 degrees. Upon closer inspection we can see that this particular arrangement puts the seven lowest layers with the strongest weighting between 10 and 45 degrees. As established in section 3.2.2, averaging over this angular range is the the same as azimuthally averaging, hence why this range of atmospheric dispersions have a much smaller range of values. This does highlight a possible limitation of our helix model. If there isn't a smooth rotation in wind direction between layers, this model could suggest a faster averaging time than is probable.

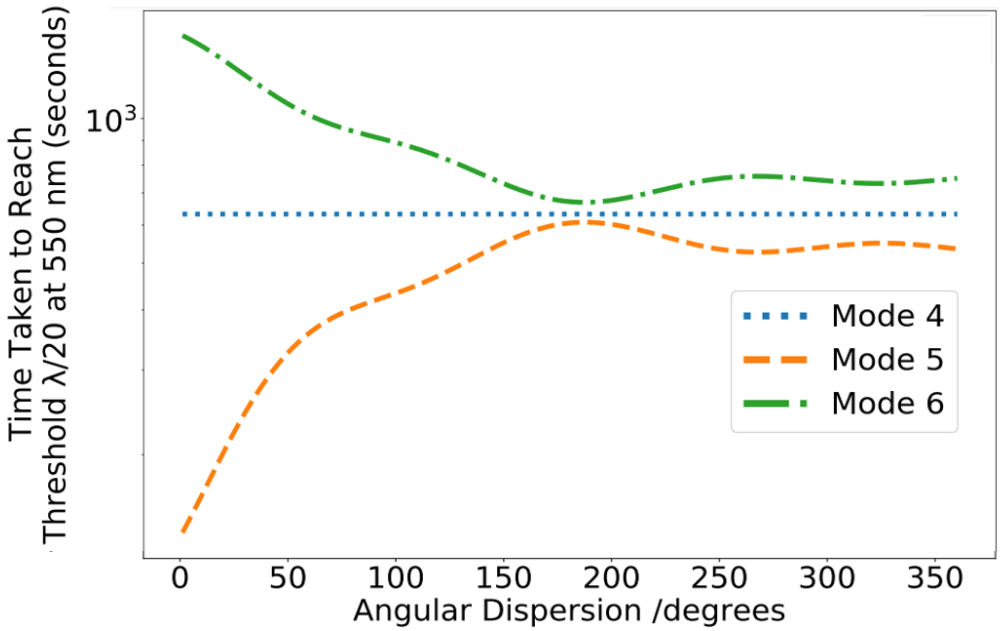


Figure 3.17: Time taken to reach an RMS WFE threshold of $\lambda/20$ at 550 nm for increasing angular angular dispersion, starting at 10 degrees. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

3.3.3 Considering Real Data

Available SCIDAR data for Paranal includes wind direction and speed for 100 different layers at variable time intervals for over 150 different nights over a 3 year period. Speed and direction data was not available for all heights as seen in figure 3.18. For this reason, for results presented here we discard SCIDAR data from above 20km in altitude, as there is less information about these high turbulent layers as they are very weak and SCIDAR algorithm tends to associate velocity values to the strongest layers. This may bias our results, but the effects should be minimal given that layers of these heights hold less than 1 percent of the total strength in our ESO profile. We show this in figure 3.19, where we have plotted the total C_n^2 value from available SCIDAR data that contains wind direction, as a function of height. There is reasonably consistent information for the ground layer, with altitudes below 1 km, and for the jet stream; which we defined as existing

between 8 km and 15 km.

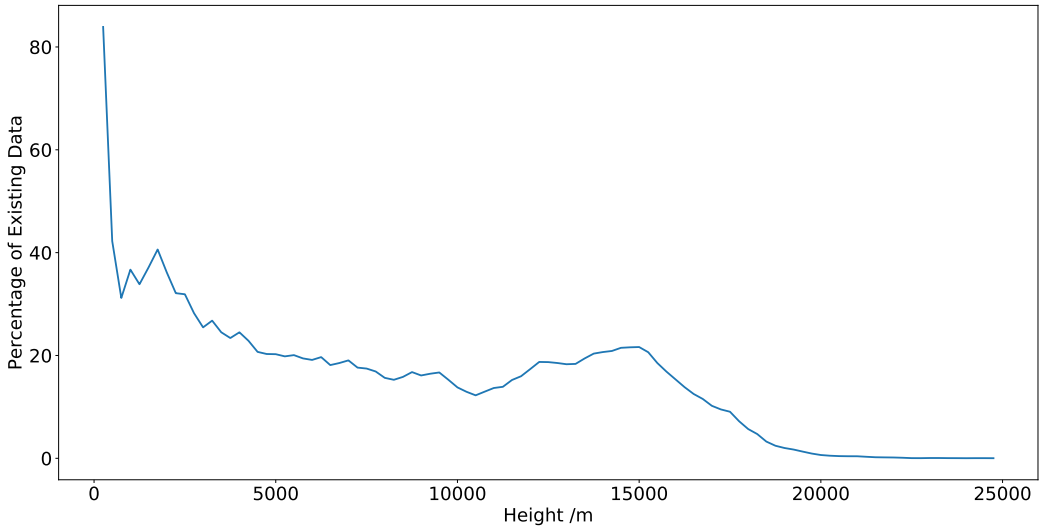


Figure 3.18: From all available SCIDAR files, across multiple nights, the percentage of data that included wind direction information, as a function of layer height.

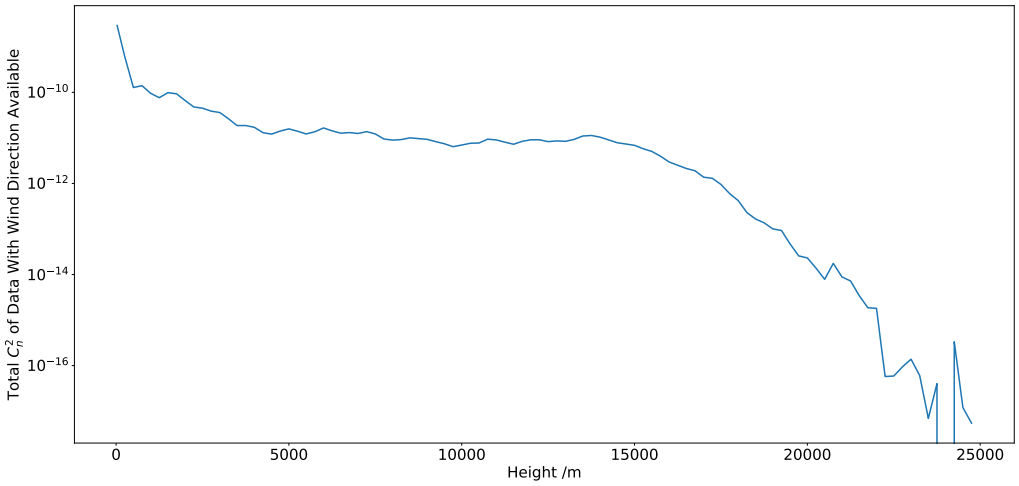


Figure 3.19: Total C_n^2 value from available SCIDAR data that included information on wind direction.

We have calculated the angular change between temporally consecutive measurements, regardless of direction. This ensures that a change between 355 and 10 degrees registers the same as between 10 and 25 degrees. Figure 3.20 shows the mean angular change between measurements on each individual night. The ground

layer tends to behave much more erratically, typically having a much wider angular change in the same time frame.

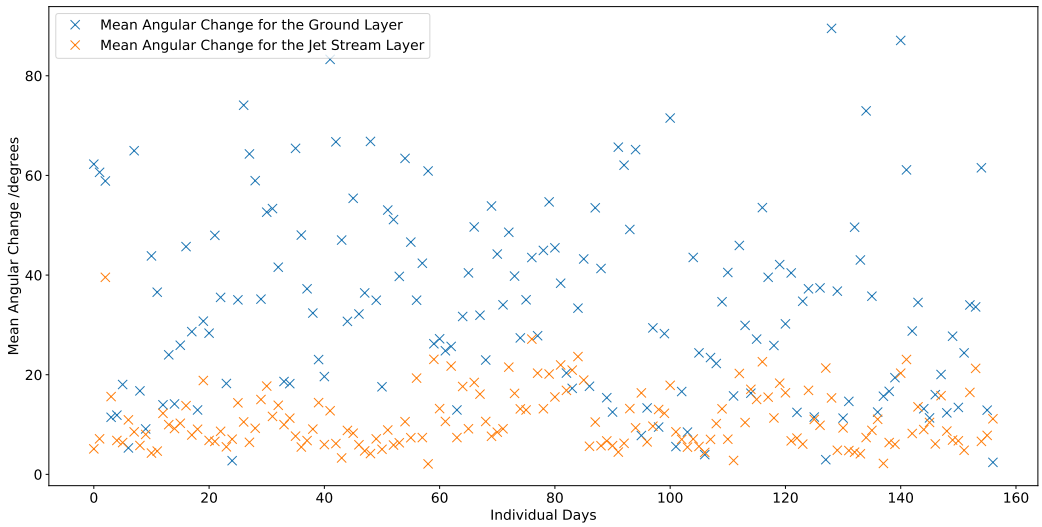


Figure 3.20: Mean angular change between SCIDAR measurements for each night of measurements; for the ground layer and the jet stream.

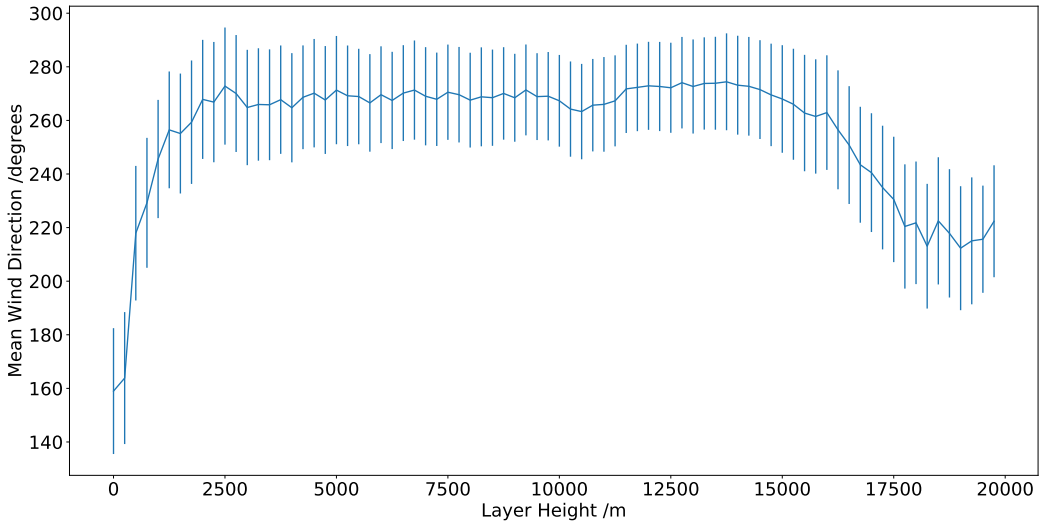


Figure 3.21: Mean wind direction with height of the layers across the available SCIDAR data. Errorbars are the standard deviation of the angular change between layers.

As well as changes in wind direction we have also shown the mean wind direction over the different measurements as a function of layer height. The average wind direction of each layer settles around 270 degrees at altitudes above 2.5km and up

to 15km, seen in figure 3.21, which is consistent with a westerly wind as discussed in section 2.1.2.1. The error bars are generated from the standard deviation of the angular change between layers at each time point. Using the mean directions from figure 3.21, we have generated a new wind velocity profile to use in our theoretical model. Each layer has a direction generated from a normal distribution with mean and standard deviation as shown in figure 3.21.

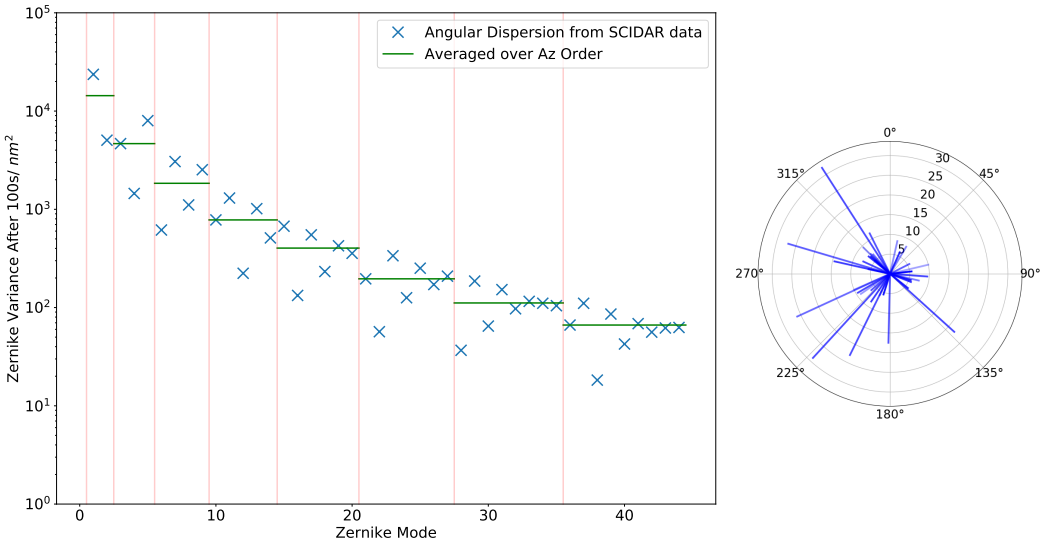


Figure 3.22: Modal breakdown of residual wavefront variance after integration time of 100 seconds for layers following a Gaussian distribution generated from the SCIDAR data. The azimuthally averaged variances are shown as the green lines. The polar plot shows the layers velocity distributions in ms^{-1}, the length of each line represents that layer's speed. Generated using the ESO 35 layer profile with $L_0 = 39$ m and r_0 of 0.157 m for an ELT sized aperture of diameter 39 m.

The results presented in figure 3.22 are consistent with others presented in this chapter. The ground layers are far from 0 degrees, so all modes are less scattered than shown in figure 3.11, but not as smoothly averaged as results that used a 360 degree helix such as figure 3.13. Higher azimuthal order modes are very close to the azimuthal average, but the WFE from tip and tilt remains an order of magnitude different to one another.

3.4 Varying Ground Layer Wind Speed

We define the ground layer as less than 1km in height, so the lowest eight layers within the ESO profile. These hold 70 percent of the weighting of the whole profile so their behaviour strongly influences the overall results. Here we broke from the wind speed profile presented in Marchetti (2015) and varied the speed of the ground layer, leaving the higher altitude layers as set in the ESO profile. As ground speed increases, the break point in the RMS will lower and so will reach a given threshold faster, as seen in figure 3.23. We have varied the speed of the ground layer from 1 to 31 ms^{-1}, the fastest speed in the original ESO profile. At low ground wind speeds averaging times can take tens of minutes to reach a $\lambda/20$ threshold. However, the ground layers are where we have the best measurements of wind speed, and wind speed variation.

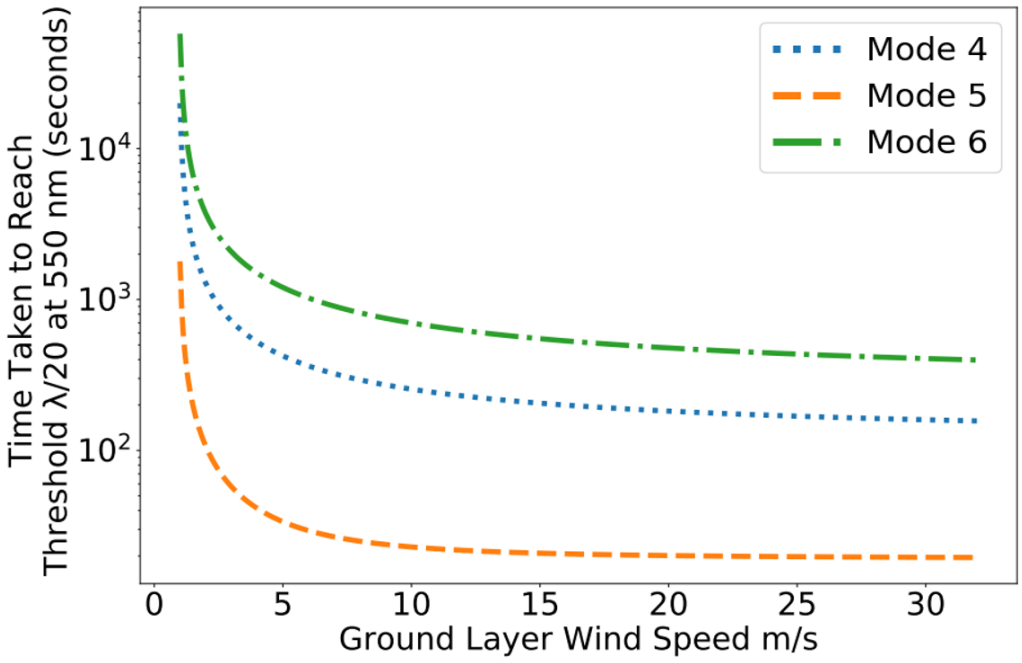


Figure 3.23: Time taken to reach an RMS threshold of $\lambda/20$ at 550 nm for ground layer wind speed in the range 1-31 ms^{-1}.

Differences between averaging times for modes within azimuthal pairs can be ex-

treme - at low ground wind speeds mode 4 can take tens of minutes to average to $\lambda/20$ whereas mode 6 takes nearly a thousand seconds.

To provide context to this we have used ground layer wind speed data from the ESO ambient conditions database ESO (1998). 2 years of data from the meteorological station at Paranal provided us with wind speeds taken at a height of 10 m, shown in figure 3.24. Wind speeds below 1 ms^{-1} have been removed as the model, which assumes Taylor frozen flow, may not be valid in this regime.

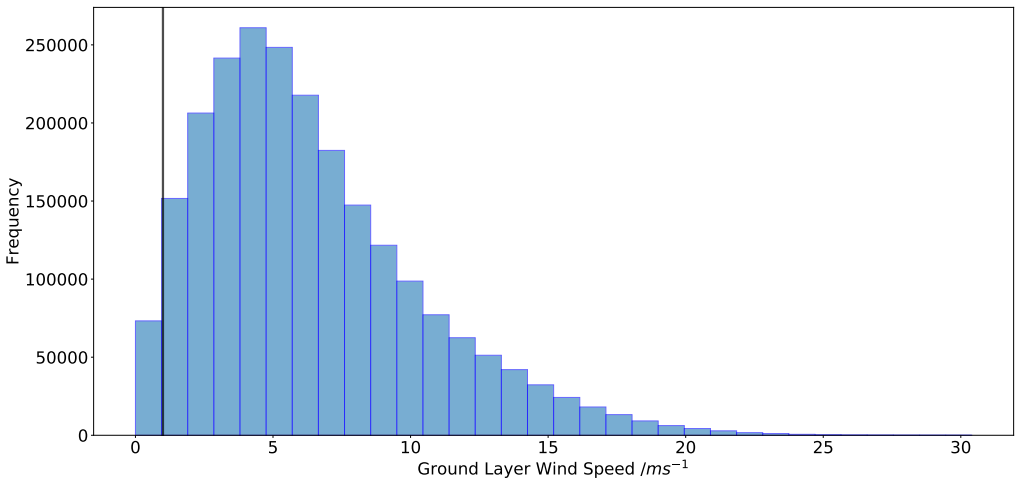


Figure 3.24: Histogram of ground layer wind speeds from the Paranal ambient conditions database. $1ms^{-1}$ are marked with a vertical line. All speeds below this have been discarded as they may not be valid within our model.

A histogram has been generated to show the frequency of times taken to reach the V-band threshold, shown in figure 3.25. For the individual modes a faster time is more likely, but each has a possible range of seconds to reach the threshold easily over an order magnitude. 10% of the time mode 5 will take an hour or longer to reach the V-band threshold, far longer than typical telescope control update rates.

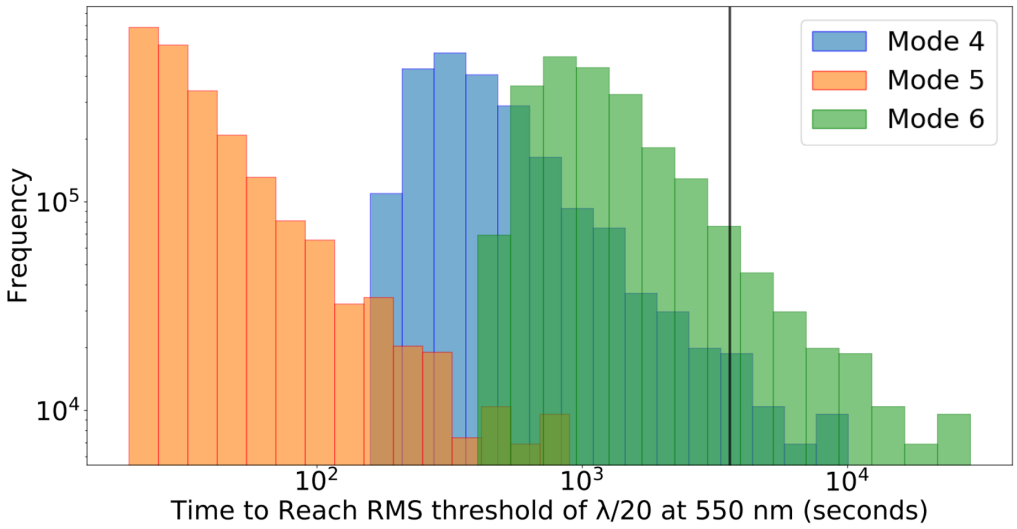


Figure 3.25: Histogram of time taken to reach an RMS threshold of $\lambda/20$ at 550 nm for modes 3, 4 and 5. 1 hour is marked with vertical line.

3.5 Conclusions

In this chapter we have tested the how considering an atmospheric wind velocity profile over a speed profile can impact the timescales for turbulence averaging on a single line of sight whilst assuming Taylor's frozen flow (Taylor, 1938). Initially in section 3.2.2 we showed that, inline with the azimuthally averaged results from Gordon et al. (2011), after a break point $\tau_{bp} \approx D/v$ the RMS for Zernike coefficients individual modes also decreases with exposure time approximately as power laws. Their gradients are dependent on both azimuthal order and the direction in which the phasescreen is moving. Wind velocity therefore can have a large impact on the time taken for a given mode to average to a specific threshold. We have shown that at ELT scales this can take thousands of seconds for low order modes.

We expanded our model to consider multiple atmospheric layers in section 3.3 and now explored the effects of the profiles angular dispersion, with both a theoretical helical variation and based on existing SCIDAR data. A larger dispersion of wind directions pushes averaging times towards the azimuthally averaged values shown in Gordon et al. (2011). These times can still be longer than the typical update rates of telescope control systems Bonnet et al. (2018). Looking at the SCIDAR data we see that the wind velocity profile changes with time. A possible extension to this work would be to use a variable wind velocity profile: changing wind direction with time.

As well as the degree of dispersion, the specific directions of lower layers have large impacts on the overall results given their relative strength. Section 3.4 varied the wind speed of the ground ground layer within the bound of existing Paranal data. This can have a significant impact on averaging time; for example changing wind speeds from 18 to 1 m/s can increase averaging time by a factor of 50.

We have expanded this work in chapter 5, by testing the impacts of wind dispersion and exposure time on an active optics system using an off-axis guide star. To create

this simulation we have used a covariance matrix library, the development of which is detailed in chapter 4.

Chapter 4

Generating A Covariance Library for Two Lines of Sight

4.1 Introduction

Two lines of sight are a common requirement in AcO simulation, for example when considering a layout of one on-axis point of interest and one off-axis guide star, seen in figure 4.1. For an individual line of sight, such as in Chapter 3, simulating varying wind directions can be achieved by applying a rotation matrix to the Zernike modes. Now when considering two pupils on a single phase screen, the separation between the pupils must remain the same with respect to the Zernike modes.

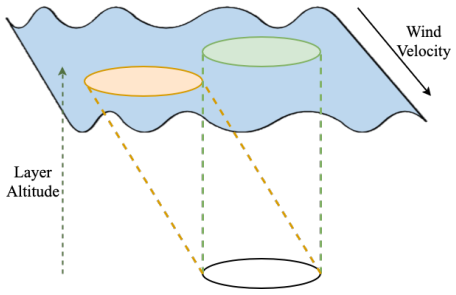


Figure 4.1: An example layout of one on-axis line of sight in green and one off-axis line of sight in orange. At a turbulent layer shown in blue, the two projected pupils are separated. This demonstrates a case in which simulating two pupils simultaneously is needed.

In section 4.2 we test two different methods to generate Zernike coefficients from

two pupils with variable phase direction and verify that the use of a covariance matrix library is both viable and useful as it saves computational space and overall computing time. The parameter spaces we chose for our Zernike covariance library are presented in 4.3.

4.2 Choice of Coordinate System

Traditionally simulating multiple pupils simultaneously required one large phase screen where, assuming Taylor's frozen flow, the pupils are moved together across the phase screen to mimic the phase blowing past. The size of the phase screen required increases in proportion to the exposure time. Simulating different phase directions required changing the direction in which they move across the screen. This is computationally intensive as each set up requires new Zernike coefficients to be calculated from a phase screen using a large amount of internal storage.

The use of an infinite phase screen reduced the size of the phase screen necessary to simulate a long exposure time as discussed in section 2.1.1 - but as new rows are only generated in one direction, the trade off is that the pupil positions have to be changed and Zernikes recalculated for each wind velocity needed. Varying wind directions whilst simulating two lines of sight through our 35 layer profile with this classical method could have become unwieldy: each layer at a different height creating a different pupil separation, for each of which Zernike coefficients would have to be recalculated for every different wind direction tested.

As an alternative, we proposed first calculating the Zernike coefficients and then varying the wind direction by applying a rotation transformation to the resultant covariance matrix. Whilst the implementation of this method is more complex, it allows one set of Zernike coefficients to represent multiple wind directions; making it on balance a more efficient system.

In this section we investigated these two methods and verify that they generate the same results:

1. The classical frozen flow method: We move the phase past the pupils at an angle, which requires a large phase screen and therefore higher computing time and storage.

2. Applying Coordinate Rotation to the Covariance Matrix (CRCM): We find and apply the transformation after calculating the time averaged Zernike coefficients.

The classical method requires running the simulation repeatedly for each wind direction we wished to study and for each variation on the lines of sight. This method is simple to understand and set up but overall computational time remains long, and storage requirements are higher given the repeated generation of large phase screens. We propose that instead we find the covariance of the Zernike coefficients and then apply the rotation directly to the coordinate system of the covariance matrix. Whilst the set up of this method is more complex, it allows one set of Zernike coefficients to represent multiple wind directions; making it on balance a more efficient system. Then in turn, a grid of Zernike coefficient covariance matrices of different pupil separations could act as a library or lookup table.

4.2.1 Pupil Geometry

To prove this CRCM method to be functional, we have shown that it produces the same covariance matrix of the coefficients as "rotating" the motion of the phase screen. Their aperture layouts on a phase screen are described in figure 4.2.

In order to ensure the cases all see the same phase screen movement in between time-steps, the layout of this test was based on a Pythagorean triple. All results should mimic phase screen motion of $-53°$, relative to Z2 (tip). The separation of the apertures in terms of pixels was divisible by 5 to avoid interpolation effects from sub-pixel offsets. We defined the positions of the apertures given the length of the baseline and the coordinates of the midpoint of the baseline. Given that the

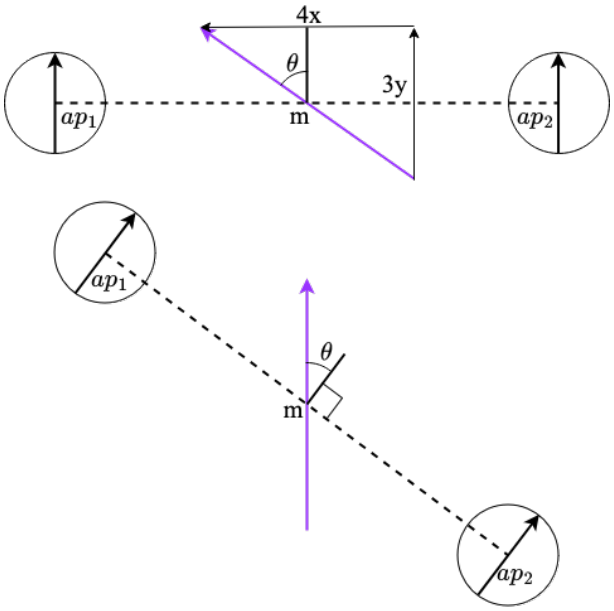


Figure 4.2: Upper: Classical method geometry of pupil separation on phase screen showing a flat baseline with angled phase screen motion. Lower: CRCM test case geometry of pupil separation on phase screen showing an angled baseline. The purple arrows show phase screen motion and the black arrows show Zernike mode orientation.

baseline = 5a and the midpoint m, the coordinates of the two apertures can be defined as:

$$ap_1 = (m - 1.5a, m + 2a) \tag{4.1}$$

$$ap_2 = (m + 1.5a, m - 2a) \tag{4.2}$$

The classical method used a baseline between the apertures perpendicular to the edges of the phase screen and angled phase motion. When using the infinite phase screens function of AOTools, new rows appear at y=0, so the phase moves vertically upwards. At the same time, we moved the apertures in a positive x direction, making the phase appear to also move in a negative x direction. By basing this motion on the Pythagorean triple triangle, the phase motion per step remained an integer number of pixels, 5a, and in the same apparent direction as seen by CRCM once the correct rotation angle has been applied.

4.2.2 Generating a Rotation Matrix

In order to utilise the CRCM method we need a rotation matrix for a series of Zernike modes. To calculate this we first derive equations for rotated Zernike modes expressed in terms of the Noll Zernikes. Starting with Noll definitions of the Zernike modes:

$$\left.\begin{aligned} Z_{even,j} &= \sqrt{n+1}R_n^m\sqrt{2}\cos m\theta \\ Z_{odd,j} &= \sqrt{n+1}R_n^m\sqrt{2}\sin m\theta \end{aligned}\right\} m \neq 0 \tag{4.3}$$

We define Z' as a set of Zernike modes that have been rotated by angle θ_0. These rotated modes can be expressed by using the Noll definitions and replacing θ with $(\theta - \theta_0)$:

$$\begin{aligned} Z'_{evenj} &= A\cos m(\theta-\theta_0) \\ &= A\big[\cos m\theta\cos m\theta_0 + \sin m\theta\sin m\theta_0\big] \\ &= \cos m\theta Z_{evenj} + \sin m\theta_0 Z_{oddj}, \end{aligned} \tag{4.4}$$

where $A = \sqrt{n+1}R_n^m\sqrt{2}$. This defines Z' purely in terms of Noll Zernikes and a rotation angle. The full set of rotated Zernike modes can be expressed as:

$$m \neq 0 \begin{cases} Z'_{evenj} = \cos m\theta Z_{evenj} + \sin m\theta_0 Z_{oddj} \\ Z'_{oddj} = \cos m\theta Z_{oddj} - \sin m\theta_0 Z_{evenj} \end{cases} \tag{4.5}$$

$$m = 0 \begin{cases} Z'_j = Z_j \end{cases} \tag{4.6}$$

This can then be expressed as a rotation matrix for a series of Zernike modes by working in pairs of coefficients, using a 2x2 block for odd or even Zernike pairs. Modes with zero azimuthal order are not rotated, so the rotation matrix is 1 on diagonal. An example matrix for the first 9 Zernike modes:

$$\begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & \cos m\theta & -\sin m\theta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sin m\theta & \cos m\theta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos m\theta & \sin m\theta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sin m\theta & \cos m\theta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cos m\theta & \sin m\theta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\sin m\theta & \cos m\theta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cos m\theta & \sin m\theta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sin m\theta & \cos m\theta \end{Bmatrix} \tag{4.7}$$

4.2.3 Comparison Between Classical and CRCM Methods

Theoretically these cases are equivalent and should yield identical covariance matrices. However they of course have small variations caused by small statistical differences across the generated phase screens and the imperfect nature of expressing circular Zernikes on a square array. We calculated the Zernike coefficients for 500 points of phase motion - limited by the size of the required phase screen for angled phase motion. This was repeated for 250 different phase screens, limited by both computing time and memory space. This gave over 100,000 coefficients for each mode to calculate the covariance.

We tested these comparisons for 4 different pupil diameter samplings - 32, 64, 96 and 128 array elements - but constant physical size. We have also set $r_0 = L_0 = D$ to remain dimensionless. This allows us to probe the possible limitations of these comparisons given the coarseness of the circular modes on square pixels. In the following sections we investigate several metrics to highlight differences between the results of the classical and CRCM methods

4.2.3.1 Instantaneous Zernike Variance Verification

The instantaneous Zernike variance from one aperture should have the same step like pattern as the Noll variances as seen in figure 2.3, but not identical given the finite outer scale. These should match across the cases if the phase screens had been generated with the same parameters, which is confirmed in figure 4.3. As expected, the approach is valid as both methods produce similar instantaneous Zernike variance values.

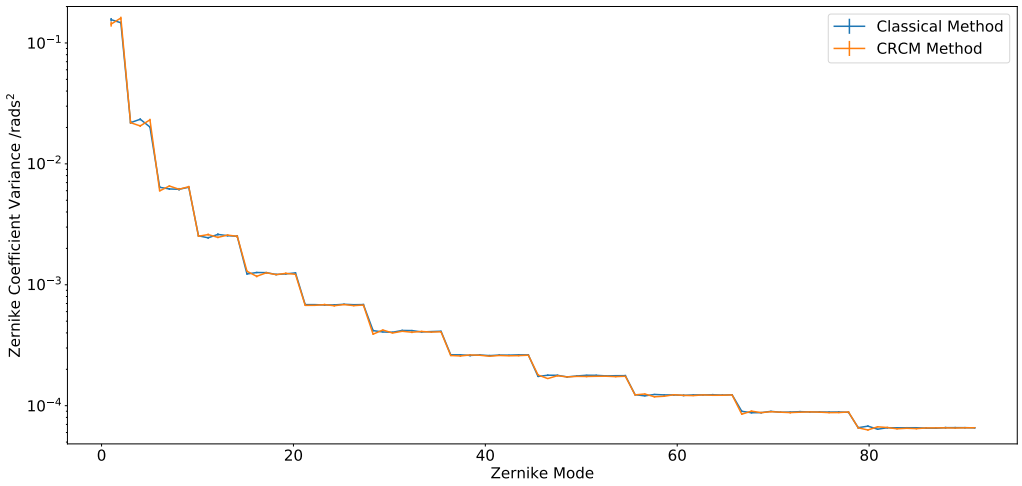


Figure 4.3: Instantaneous Zernike Variance produced using the CRCM and the classical method for a pupil sampling of 96 pixels.

4.2.3.2 Difference Matrices

To compare the similarities in two different covariance matrices, a normalised difference matrix was used,

$$\Delta_{ab} = (C_a - C_b)/C_{rmsa}, \tag{4.8}$$

where the difference between two covariance matrices C_a and C_b is normalised by the RMS of the first:

$$C_{rmsa} = \sqrt{\langle \mathcal{C}_a^{sub^2} \rangle} \tag{4.9}$$

This RMS was calculated by splitting the coefficients into 10 sections of 12,500, then finding the covariance matrix, $\mathcal{C}_a^{sub}$, for each group. Subsequently the RMS

across these 10 covariance matrices is found. This allowed us to more clearly see the difference between two matrices that includes values close to and either side of 0. We then split Δ_{ab} into different radial orders as shown in figure 4.4.

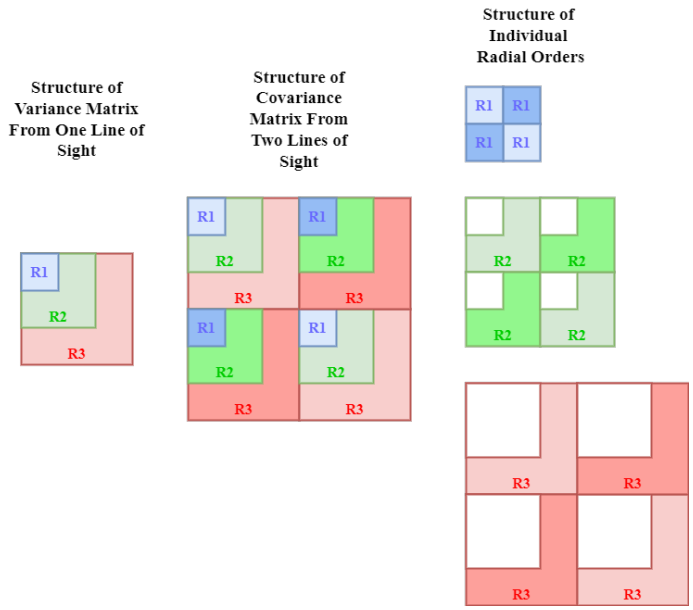


Figure 4.4: From left to right: Structure of a variance matrix from a single line of sight containing three radial orders. Structure of covariance matrix from two lines of sight containing three radial orders. Structures of each individual radial order used when showing the maximum difference values.

We used the first 12 radial orders for this test, so each comparison resulted in 12 variably sized difference matrices. To show these results graphically we present the maximum value in Δ_{ab} from each radial order.

4.2.3.3 Measurement of Statistical Noise

The comparisons between the classical method and our CRCM method are not expected to be completely identical due to limited sampling. Calculating the difference between two covariance matrices both produced with the classical method of phase rotation or two cases using CRCM will establish a criterion for "agreement". We see that the differences between repeat cases remain reasonably constant with variation in pupil sampling.

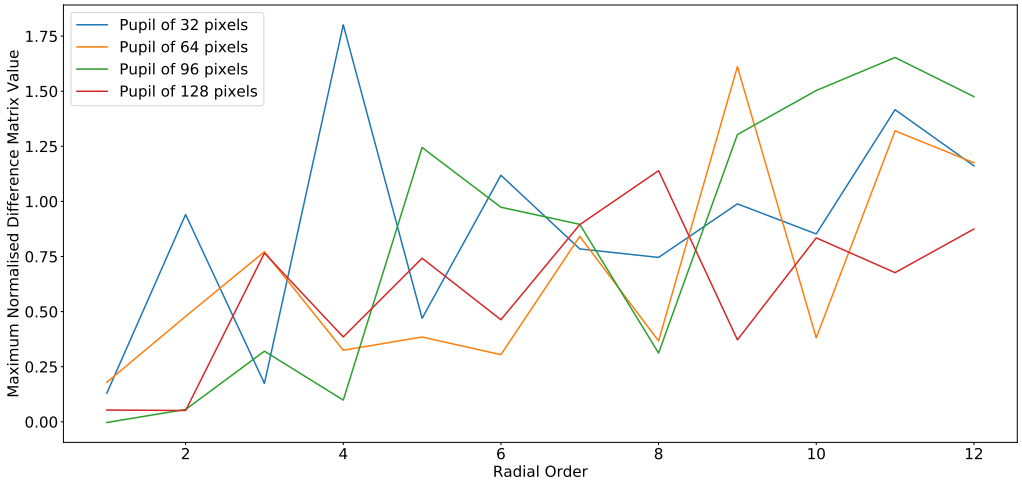


Figure 4.5: Maximum difference value across radial orders of the difference matrix between two repeated cases of the CRCM method for pupil diameter of 32, 64, 96 and 128 pixels.

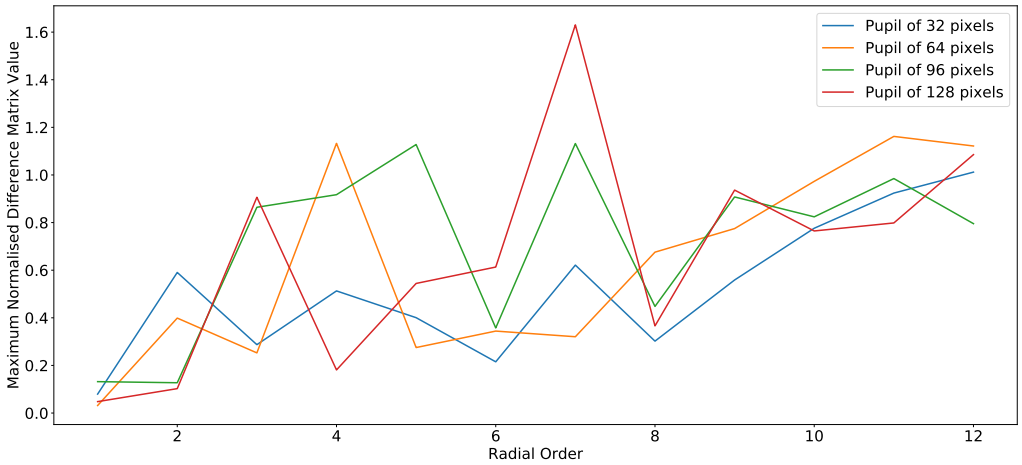


Figure 4.6: Maximum difference value across radial orders of the difference matrix between two repeated cases of classical rotation method for pupil diameter of 32, 64, 96 and 128 pixels.

Figures 4.5 and 4.6 show a slight increase in difference for lower order modes when using the smallest pupil sizes, but results remain of a similar scale for all four scales. These covariance matrices were produced with identical methods but independent phase screens. These results show that there is a non-zero difference, which appears to be more pronounced for higher radial order Zernike modes. This statistical noise could be attributed to the number of coefficients used to create the

covariance matrices or the phase screens. These are a benchmark of comparability, the difference matrices between the classical and CRCM methods should be no larger than the peak value of 1.75 shown in figure 4.5.

4.2.3.4 Comparison between Mismatched Cases

As an example of a poor equivalence between two covariance matrices we have compared one case using the same CRCM method and pupil orientation as above against a purposefully mismatched scenario. This used the classical method with a wider separation of 8a pixels. For the classical method the only phase motion came from the addition of new rows from the AOTools infinite phase screen. This led to a phase angle difference of +53 degrees between the two cases. The separation 8a was chosen as the largest baseline that would fit in a phase screen of the same size used for the other tests.

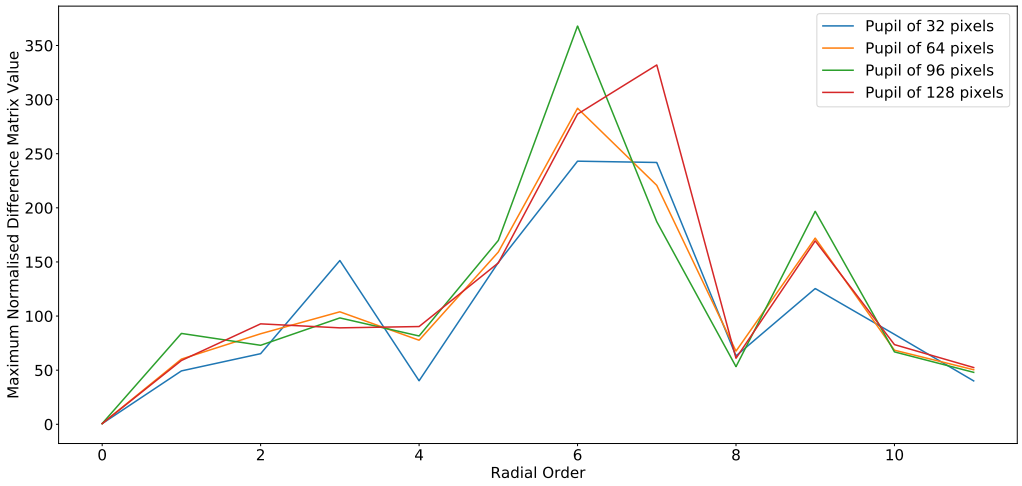


Figure 4.7: Maximum difference value across radial orders between mismatched pupil separations and apparent phase rotation. Common parameters between simulations are the averaging distance and that $r_0 = L_0 = D$. This is shown for pupil diameters of 32, 64, 96 and 128 pixels.

As expected, figure 4.7 shows a consistent increase in difference for all radial orders and all pupil sampling. The maximum differences from figure 4.7 are at least an order of magnitude larger than the repeated test cases shown in section 4.2.3.3.

Taking the 64 pixel sampling as an example, for radial order 12 the difference has jumped from ∼1 to just over 350. This logic is reasonable when we consider how heavily the averaged results were effected by wind direction in Chapter 3, it follows that the whole covariance matrix will also be sensitive to changes in phase direction.

4.2.3.5 Cross Verification Between Classical and CRCM Methods

We now calculate the difference matrix between the two different coordinate systems with the same angle of phase motion relative to both the Zernike modes and the pupil orientation.

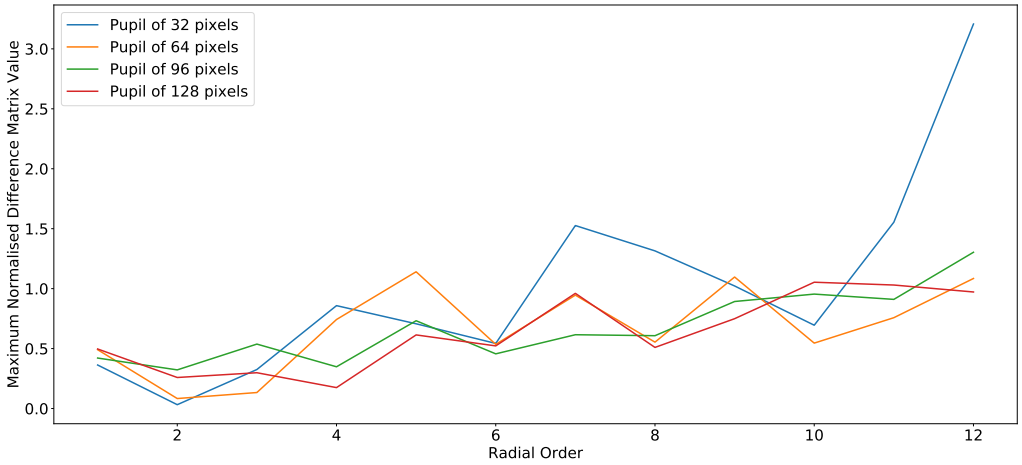


Figure 4.8: Maximum value across radial orders of the difference matrix between the classical method of phase rotation and using the CRCM, for pupil diameter of 32, 64, 96 and 128 pixels.

The maximum values from the difference matrix of this comparison are shown in figure 4.8. This does show a slight change in the difference matrix for pupil sampling of 32 pixels. For radial orders 9 through 12 the difference value increases from 0.7 through to 3.2, whereas the larger pupil samplings remain constant in scale. We had expected to consistently see that coarser sampling caused a larger difference for the higher radial orders across sections 4.2.3.3, 4.2.3.4 and 4.2.3.5. As the higher order modes have steeper gradients, coarser sampling would struggle to

accurately represent a more complex circular pattern on a square array. It would seem however, that the use of the Gram-Schmidt orthonormalisation process has reduced this effect, except in figure 4.8. Higher order modes still have a slightly higher difference when comparing between any two covariance matrices - from ~ 0.5 to $\sim 1\text{rad}^{-2}$, but this is not noticeably affected by sampling. When moving forward with our covariance library, we have still limited our covariance matrices to the first 10 radial orders of Zernike modes. Pupil sampling of 96 pixels has been chosen for historical reasons - due to initially performing these tests prior to applying the Gram Schmidt process.

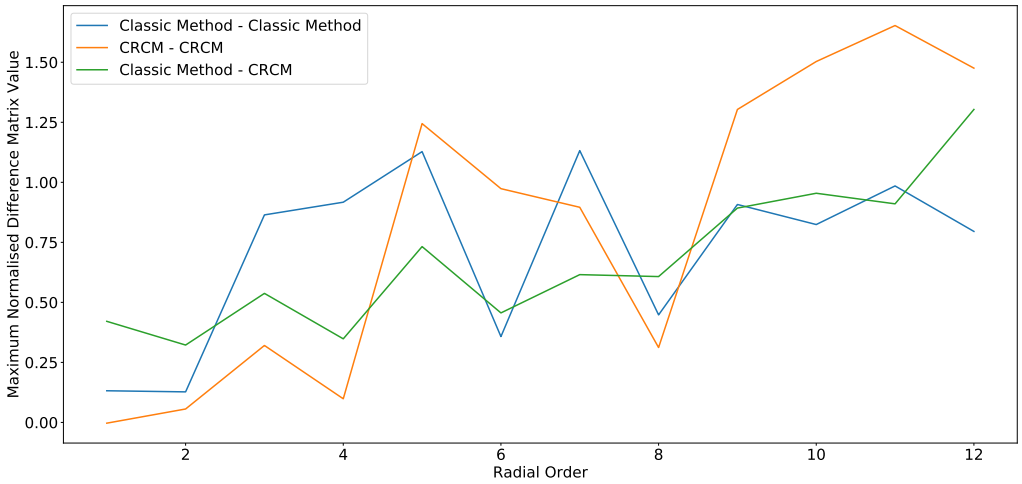


Figure 4.9: Maximum value across radial orders of the difference matrix for pupil diameter of 96 pixels. Shown are the difference values between two cases of both the classical and CRCM methods and a comparison between the classical and CRCM methods

Figure 4.9 shows three maximum difference values, all with pupil sampling of 96 array elements in diameter: 2 independent runs of the classical method, 2 independent runs of the CRCM method and a comparison between one independent run of each. All three show a slight increase in difference with radial order but are all of comparable scale. The comparison between the two methods is larger than the repeating comparisons for radial orders 1, 2 and 6, but only reaches 78% of the maximum 1.55 value of the comparison between two cases of using CRCM at this pupil sampling. This was sufficient evidence to say that the comparison

between using a coordinate transformation of the covariance matrix and the classical method of rotating the phase motion were the same within statistical error; and the two methods are equivalent.

4.3 Defining the Parameter Space for a Covariance Library

Having established that applying a coordinate transformation to a Zernike coefficient covariance matrix between two pupils to represent different wind velocities is a valid method, we set out to generate a library of covariance matrices that represent a range of pupil separations and averaging times. This series of pupil geometries with a fixed wind direction could then be re-scaled and rotated for various wind velocities. In this next section we discuss the limits of this method and the justifications for the parameter spaces we chose.

4.3.1 Method for Calculating Covariance Matrices

In this section we discuss the method used to create a covariance library from an individual phase screen. A grid of Zernike covariance matrices first required a grid of Zernike coefficients. A row of n_x pupils were positioned perpendicular to phase screen motion where we measured n_y instantaneous Zernike coefficients, again utilising the infinite phase screen, where $n_x << n_y$. Using Taylor's frozen flow approximation, a spatial offset in the direction of phase screen translation is equivalent to a temporal offset; meaning later coefficients for this row of pupils were equivalent to an earlier measurement further along the phase screen. This meant the large number of Zernikes from the initial row could be cut into shorter series of coefficients for an apparent grid of pupils, meaning only one infinite phase screen was necessary rather than repeatedly generating new ones. This method is

demonstrated in figure 4.10. This produced a spatial grid of pupils, each with a long series of instantaneous Zernike coefficients.

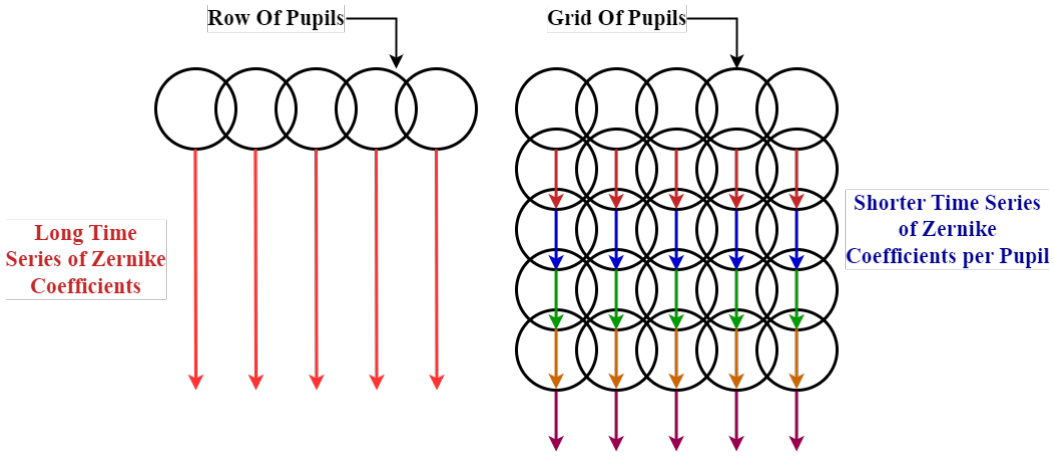


Figure 4.10: Left: As the phase screen moves the row of pupils calculate a long series of Zernike coefficients. Right: The long list of coefficients from the row of pupils becomes a grid of pupils with smaller lists of coefficients.

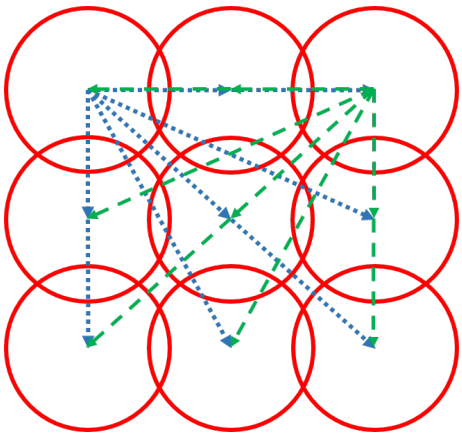


Figure 4.11: Ensuring all baselines are covered, here for a smaller grid example. The first and last pupil in the first row is compared to every other pupil in the grid.

To determine all possible baselines required for the covariance matrices, only the first and last pupil from the initial row are compared with all others in the grid as shown in figure 4.11, making the total pairs of pupils - and therefore files - $2(pupil\,row)^2$. The total number of pupils within the grid is dependent upon the spatial sampling of the initial row of pupils, and its length. This meant the covariance library covered pupil separations in all directions for separations, but only

within the bounds of the total grid size. Pupil separations with non-integer pixel spacings were reproduced by interpolating between points in the covariance grid. A grid of covariance matrices has been generated for three different D/L_0 ratios: 1, 3 and 10, to allow the library to produce results for a range of telescope diameter sizes.

4.3.2 Setting Parameters

The covariance library must contain enough detail to be useful for further calculations, but not unnecessarily detailed to detract from any possible improvement in computing time. In this section we have tested and constrained the separations between points in the grid in terms of both spatial, δ_x and δ_y, and temporal separations T_{av} in order to make the library as small as possible without introducing significant errors. We have also set the maximum size of the grid m_x, after which the covariance values are small enough to not include in the library. We demonstrate these parameters visually in figure 4.12.

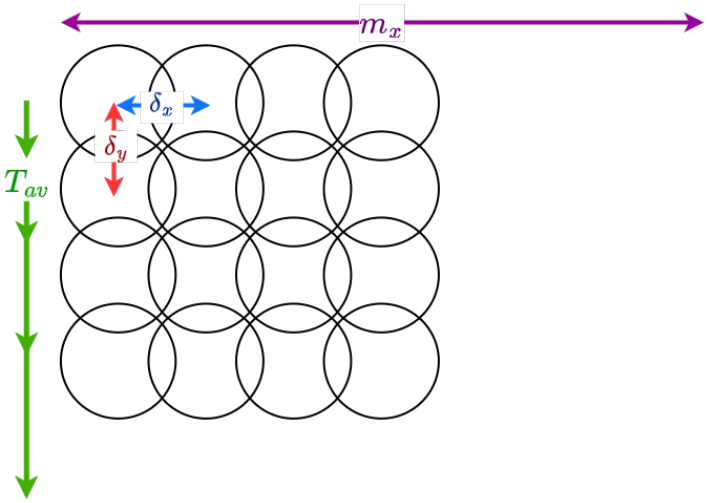


Figure 4.12: Demonstrating the constraints we put on the covariance matrix grid. The spatial separation between pupils, δ_x and δ_y, the maximum spatial width of the grid m_x and the distance between temporal averaging steps T_{av}.

4.3.2.1 Maximum Pupil Separation

We constrain the size of the covariance grid to minimise the total size of the covariance library. To do this we have used simulation to replicate the normalised spatial covariance as seen in figure 8 (Wilson and Jenkins, 1996). This shows covariance with no time averaging - normalised by zero offset variance values - against pupil separation and eventually the separation between the pupils was so large that there was no longer a correlation between the covariance of the two pupils. The maximum pupil separation can be set by the separation where the covariance is zero. *

Figure 4.13 recreates the Wilson and Jenkins covariance for modes with the same azimuthal order, showing they follow similar trends before settling around zero, but the higher radial order reaches zero faster – the limiting factor to the maximum size of the grid m_x is the lowest order modes.

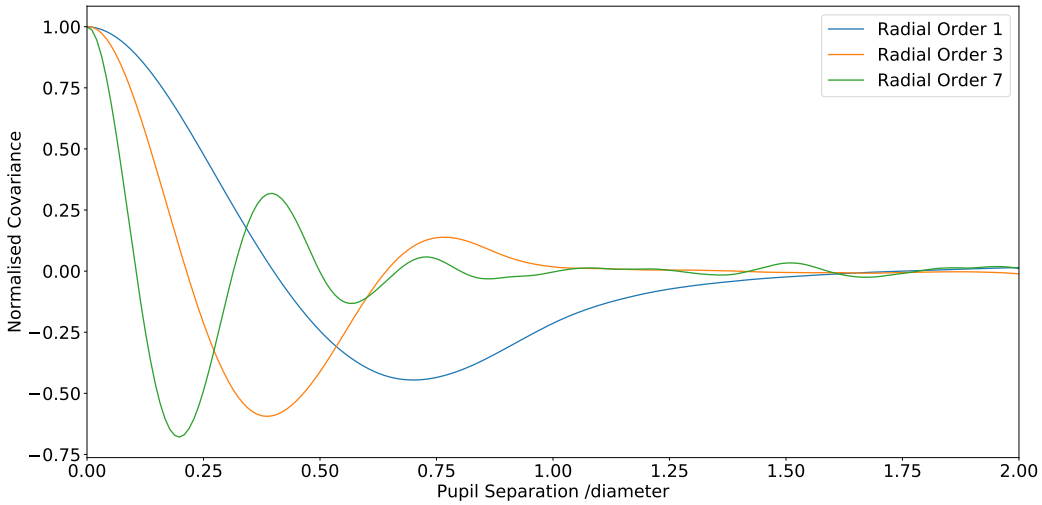


Figure 4.13: Instantaneous Zernike covariance normalised by variance as a function of pupil separation for azimuthal order 1 modes. This was generated using a single layer of turbulence with $r_0 = L_0 = D$. The normalised covariance fluctuates as the pupil separation increases before settling around zero.

Following on from this we checked the effect of outer scale on the normalised cov-

*Note that in Wilson and Jenkins (1996) there is an error in the x axis of figure 8, which should read 0-2 rather than 0-4, Figure 8 is also given in terms of pupil radius rather than pupil diameter.

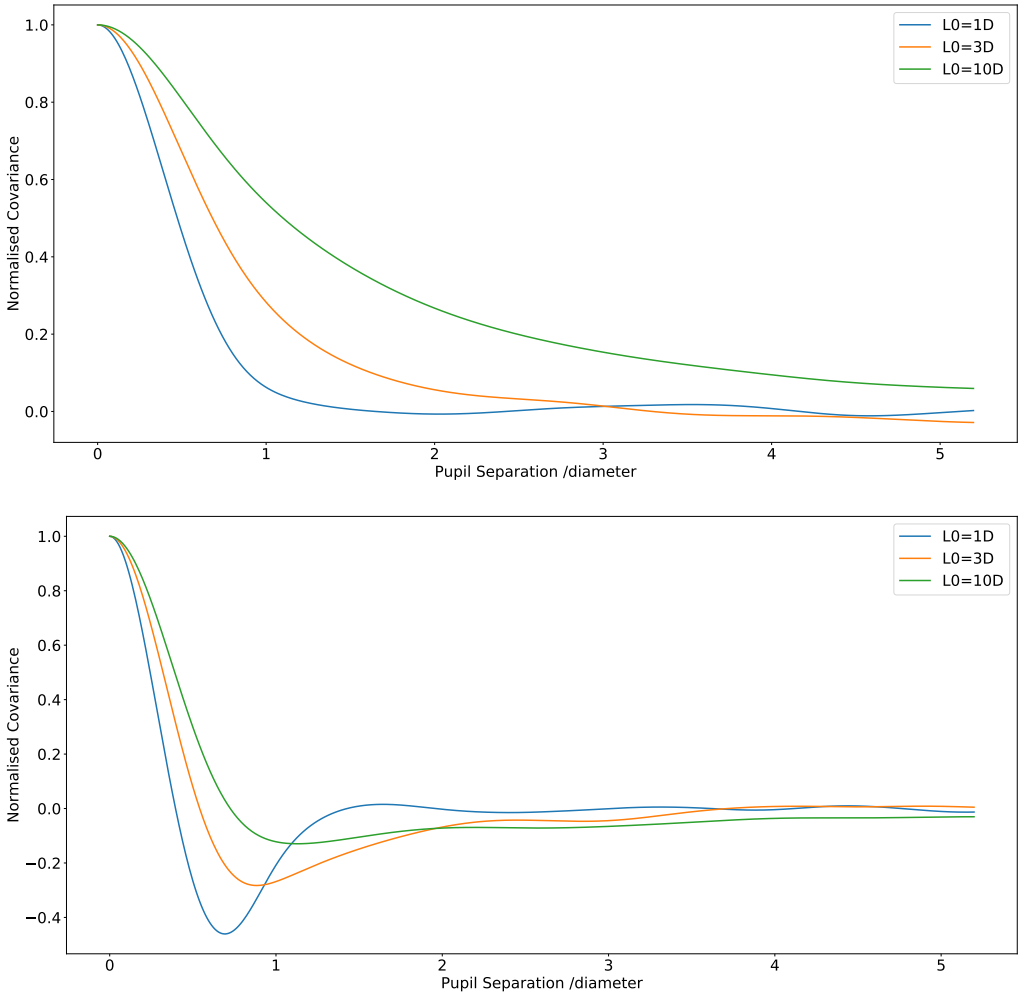


Figure 4.14: Instantaneous Zernike covariance normalised by variance as a function of pupil separation for modes tip and tilt. This was generated using a single layer of turbulence with pupil diameter $D = r_0$, with outer scales of $1, 3, 10 \times D$. The normalised covariance fluctuates as the pupil separation increases before settling around zero at larger separations for larger outer scale values relative to pupil diameter.

ariance for the lowest order modes: tip and tilt. For larger outer scales relative to pupil diameter, the Zernike covariance will reach zero at larger separations. This is shown in figure 4.14. Wavefront tilt settled around zero by $\sim$ two pupil diameters of separation for all three tested outer scales in figure 4.13. Wavefront tip however, takes much longer to reduce. The change between the two modes is due to the phase direction. If phase motion was perpendicular to this, the behaviour for tip and tilt would be reversed. m_x was set at 4 pupil diameters, meaning that tip and

tilt cannot be accurately extended off the spatial grid, but the other 9 radial orders can be.

4.3.2.2 Spatial Step Size

Having constrained m_x, the maximum spatial separation of the grid, we move to find δ_x, the step size between pupil locations within the grid. It is unlikely that all pupil separations we wish to use will fall directly on a grid point, so δ_x must be small enough to accurately interpolate to any other position in the grid space.

To find the maximum acceptable step size between pupil separations we first calculated the same normalised covariance from section 4.3.2.1, with pupil separations increasing with the minimum possible step of one pixel. From this we pulled out the equivalent of larger step sizes and judged the accuracy of interpolating back to the finer sampling.

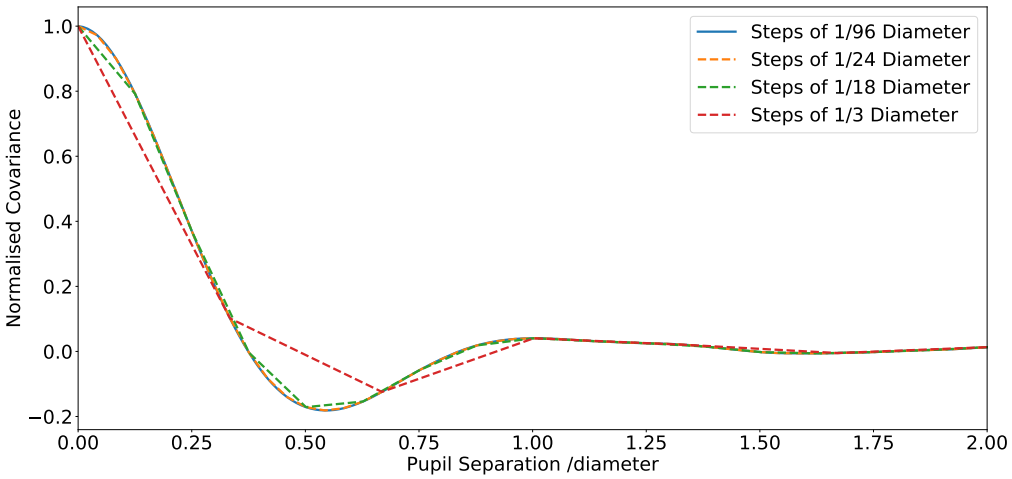


Figure 4.15: Use of linear interpolation for a range of spatial samplings for the normalised covariance of focus (mode 4).

As expected, figures 4.15 and 4.16 show the interpolated curves begin to deviate from the target values as the sampling gets coarser. The different spatial offsets are stored in different files so we have used linear interpolation here, meaning only the two files either side of the relevant point are needed. As these curves oscillate

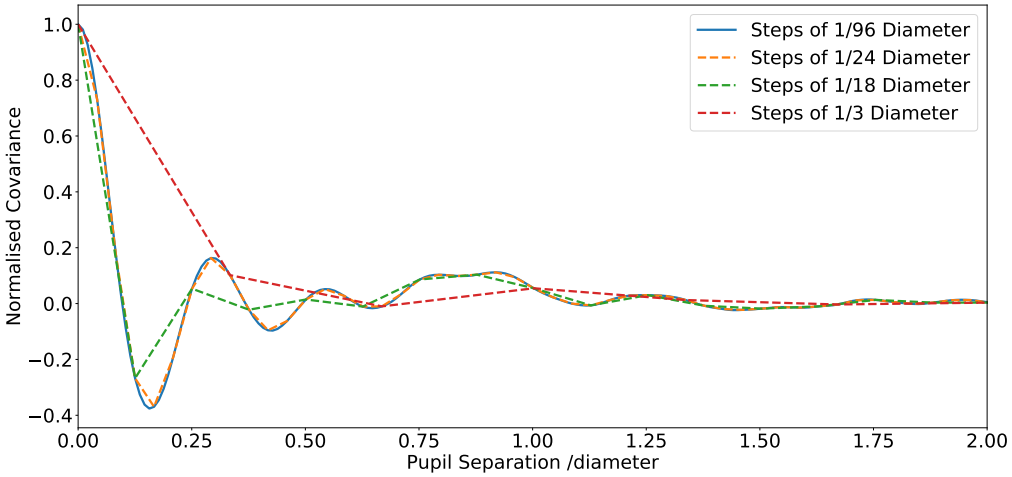


Figure 4.16: Use of linear interpolation from a range of spatial samplings for the normalised covariance of mode 66.

either side of zero, the normalised difference cannot be used to show the accuracy of the interpolation. Instead we find the RMSE for a rolling short range of pupil offsets – moving in sections of 1/3 pupil diameters to cover the initial dip shown in figure 4.16.

The rise in error for smaller pupil separations is present in all pupil steps - the use of cubic interpolation would reduce it by an order of magnitude as seen in figure 4.18, but would require loading a much larger proportion of the grid every time a desired separation did not fall exactly on a saved file, increasing overall computing time when utilising our covariance library. The RMSE for steps of 1/8 pupil diameter at large separations is of a similar scale to the error from smaller steps at smaller separations; indicating that non-linear sampling across the grid may not have had a noticeable effect in the performance of the covariance grid. Despite the fact that this could have also reduced the number of files in the grid we have elected to keep linear sampling, given that the different baselines will run in both positive and negative orientations (see figure 4.11). It was simpler to have a constant step in between pupils, rather than a non-linear series running in two directions.

Figure 4.17 shows this RMSE for linear interpolation with a range of δ_x steps.

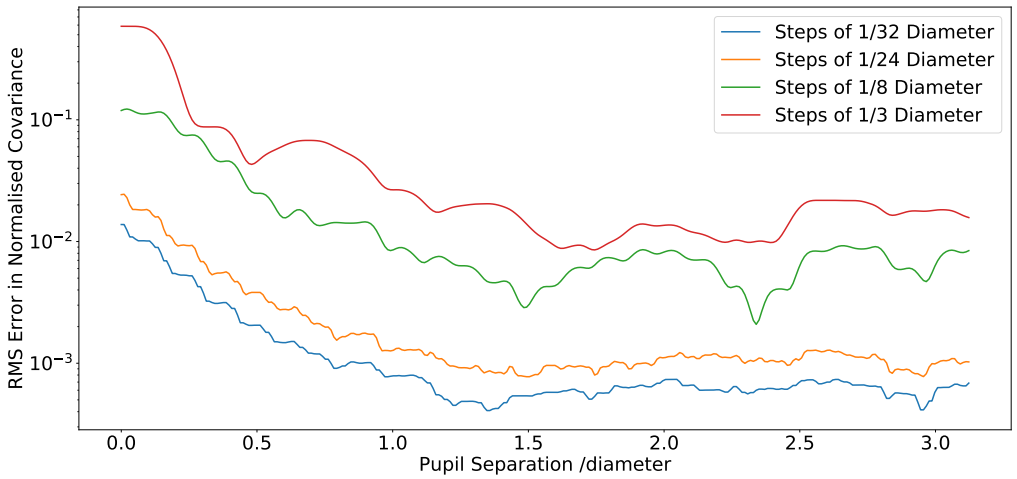


Figure 4.17: RMSE in interpolation from coarser spatial sampling of the normalised Zernike covariance of mode 66. RMS was taken from rolling sections of 1/3 pupil diameter. From this result we have set δ_x to 1/24th the pupil diameter.

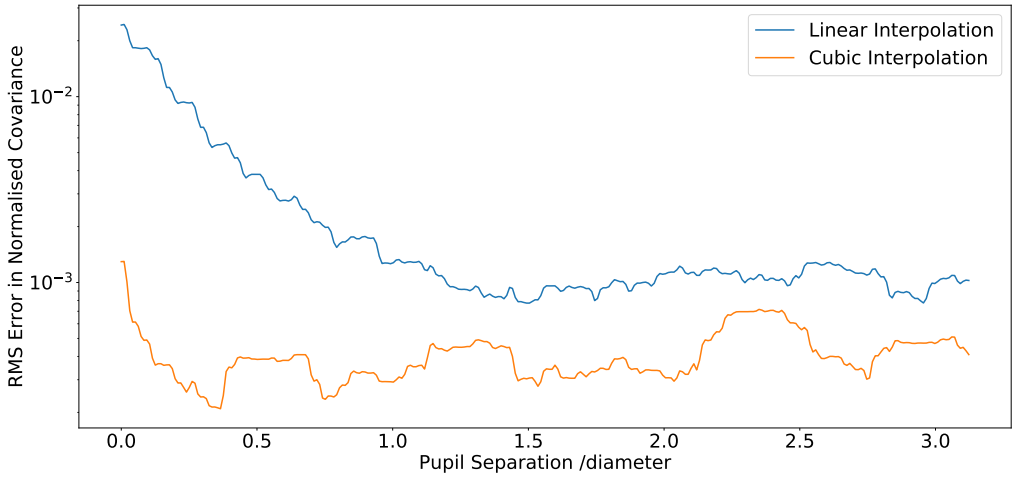


Figure 4.18: RMSE in linear and cubic interpolation from steps of 4 pixels of the normalised Zernike covariance for mode 66. RMS was taken from rolling sections of 70 pixels.

Steps of 1/32D and 1/24D pupils have a relatively similar error of 1.5 and 2.4% respectively, so we additionally consider computational storage. For the established grid width m_x of four pupils with 96 pixel diameters, steps of 3 or 4 pixels translates to 96 or 128 pupils per row respectively. This means an overall difference of over 7000 files as part of the full grid. Therefore the pupil separations δ_x and δ_y were set at 1/24 pupil diameters. When linear interpolation is used, this gives a worst

case error in the normalised covariance of ±2.4%.

4.3.2.3 Temporal Separations

Having defined the maximum spatial size of the grid and the separations of the pupils within it, in this section we define the temporal separations between the time averaged Zernike coefficients, T_{av}, used to create the different covariance matrices for each pupil baseline. The individual covariance files include matrices for the same pupil separation with increasing time averaging. The time averaging steps are therefore the major factor in the size of the individual files - and should be reduced as far as possible whilst remaining fine enough to utilise the covariance library without additional errors

As shown in figure 4.10, the spatial y-axis and time axis are the same for the translation of the screen in y. The chosen δ_y of four pixels was therefore also the maximum number of phase screen rows moved in-between calculated Zernike coefficients. Whilst the ideal here would be to keep the step size for the coefficients the same as the grid spacing, the initial coefficients have to be close enough together to identify the initially correlated Zernike RMS WFE for very small exposure times, before the break point, as seen in figure 3.1. As all the time averaging steps are saved within the same file, cubic interpolation can be used. We also note that the higher order modes have a break point for smaller averaging distances as seen in figure 3.1 - the higher order modes are the limit on the size of the T_{av} steps. For this reason we use mode 66, the last mode in radial order 10, to test T_{av} as it is the highest order mode we plan to save in the covariance library.

To remain dimensionless, we show this temporal averaging in terms of averaging distance in the units of pupil diameter. We have calculated the time averaged variance of mode 66 for increasing averaging distance up to 1 pupil diameter, increasing in the minimum steps of one pixel, shown in figure 4.19. We have then interpolated back to this result from steps of 1/48, 1/32 and 1/24 pupil diameters. We then

calculated the rolling RMSE as previously shown in 4.3.2.2, this is shown in figure 4.20.

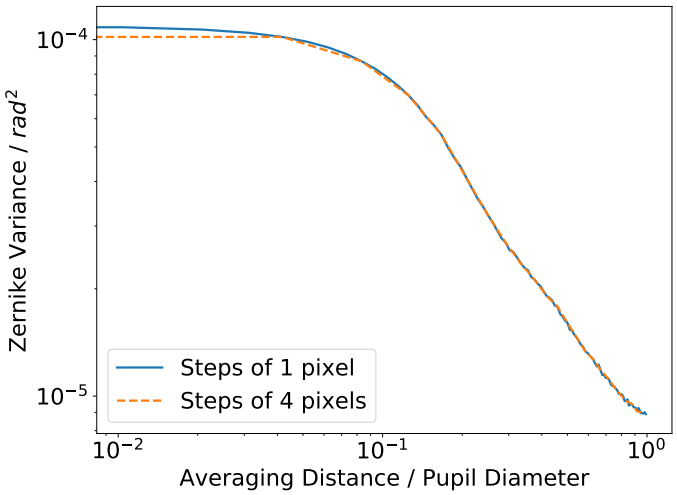


Figure 4.19: Variance of mode 66 for increasing averaging distance, measured in steps of one and four pixels.

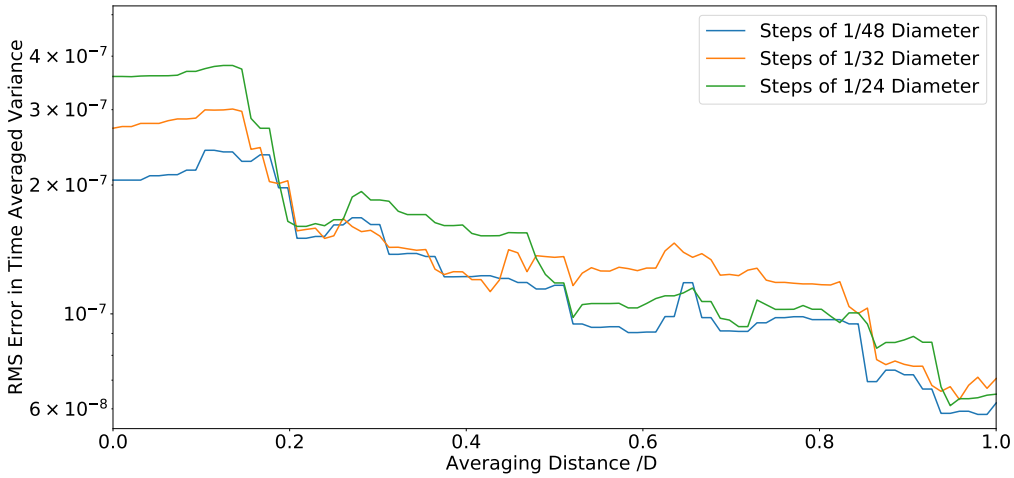


Figure 4.20: RMSE in cubic interpolation from steps of 1/48, 1/32 and 1/24 pupil diameters of the variance of mode 66 for an increasing averaging distance to one pupil diameter. The RMSE was taken from rolling sections of 5/24 pupil diameters.

Figure 4.20 shows that there is very little difference the interpolation error between moving in steps of two, three or four pixels for mode 66. Also shown in figure 4.19, is the Zernike variance of mode 66 in steps of both 1 and 4 pixels for a short averaging distance up to one pupil diameter. As it is convenient and has no noticeable effect on the error, we set T_{av} initially in steps of four pixels.

The maximum time averaging was set at 100 pupil translations, over a total series of 20,000 pupil translations. 100 pupil translations is the same as the maximum separation in the results presented in Chapter 3. For example, for a 10m diameter pupil and a wind speed of 5 ms^{-1}, the covariance library would cover 200 seconds. With 1/24 pupil diameters between each coefficient sampling point, this gives a maximum of 2400 different covariance matrices for each pair of pupils.

From figure 3.3, we see that the change in time averaged variance between averaging over 1 and 1.1 pupil translations is much more significant than the change between 99 and 99.1. We can therefore reduce the overall file size by working in non-linear averaging steps. Given this, a non-linear series of 100 T_{av} steps has been used. The non-linear steps in δ_y exponentially increase with base 2. The error in the coarser sampling is compared to the linear sampling of 1/24 pupil diameters in figure 4.21 demonstrating a consistently small error across the full separation. Whilst the non linear steps do produce a slightly larger error, it remains at most $\pm 0.8\%$ of the variance value and is not visible when plotted. The change in time averaging steps reduces the file size from over 320,000KB each to under 14,000KB. By saving as 32-bit data this reduces again to 6,180KB for each file within the covariance grid.

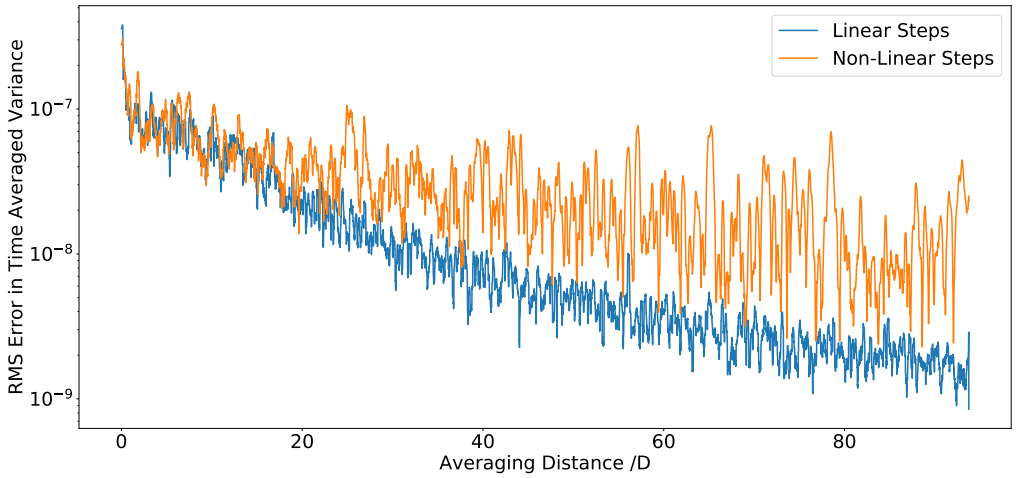


Figure 4.21: RMSE in cubic interpolation from a linear and non-linear range of averaging steps of the variance of mode 66 for an increasing averaging distance to 100 pupil diameters. RMS was taken from rolling sections of 20 pixels.

4.3.3 Going Off the Grid

There are naturally scenarios in which we wanted to use the covariance library that include either pupil separations or exposure times that are outside the modelled values of the grid. To allow for this we created a specific covariance file to represent larger separations. Off the edge of the grid we have established that we can treat the spatial covariance values as zero. The variances of the individual pupils do not depend on their position, so we take an average of all variances across the grid. If there are averaging times needed that are longer than stored in the library, then we use our power law fit as described in section 3.2.2 to extend the time axis for the variance values. This is shown graphically in figure 4.22. The extensions of the grid overlap the data from the covariance library to demonstrate the accuracy of the fit. We are only able to extend the grid for on-diagonal variance and covariance values, however these are the only ones currently necessary for our science cases.

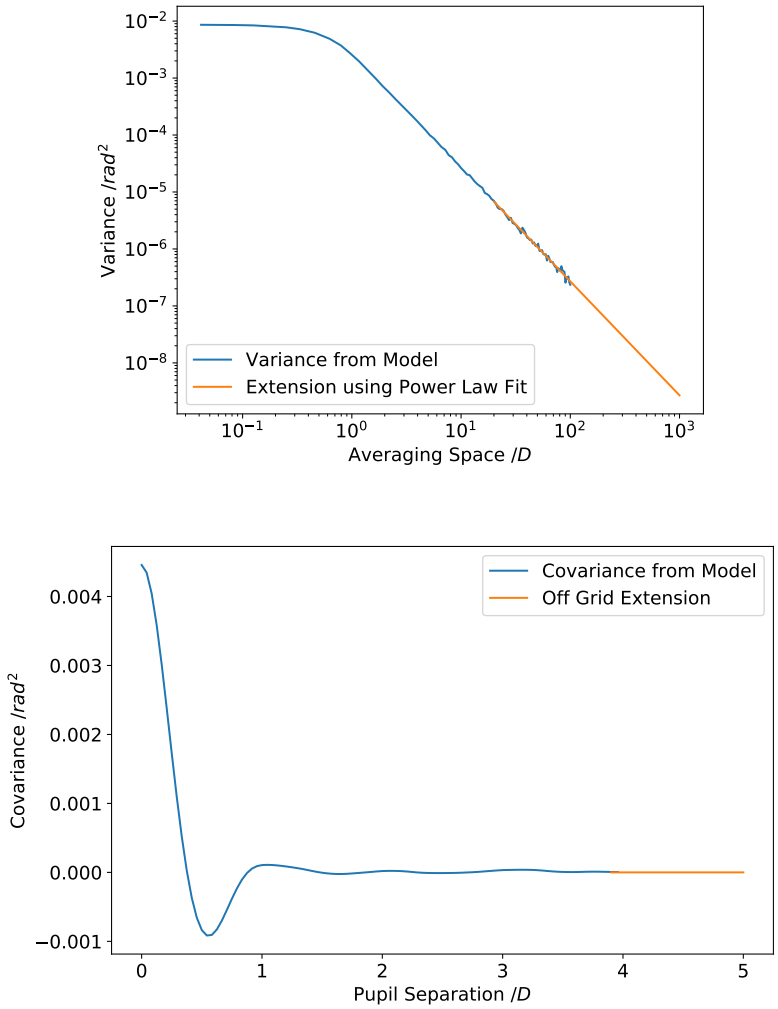


Figure 4.22: Demonstrating how we extend our variance and covariance values in both the spatial and temporal axes. Top shows the use of an approximate power law fit to increase the averaging time for the variance of mode tilt. Bottom shows that we can approximate the covariance of mode tilt between two pupils as the separation increases off the grid as zero.

4.4 Conclusions

In section 4.2 we have shown that when simulating different wind directions for two lines of sight, rotating the Zernike modes and rotating the phase motion produces an equivalent covariance matrix to applying the rotation directly to the covariance matrix. Much of the work presented in this chapter has been investigating how the choice of simulation parameters and approximations affect the determination of covariance whilst minimising the size of the library.

4.4.1 Defining Covariance Library Parameters

We have chosen and justified appropriate parameters for the covariance files, their separation and the overall size of the grid.

- Number of Zernike modes, fixed at 66 modes, the first 10 radial orders.
- Pupil diameter in pixels, fixed at 96 pixels.
- Pupil diameter and $r0$. These remain constant at $D = r_0$, so that the covariance matrices can be re-scaled when used.
- T_{av} Number of Zernike coefficients to be averaged over, up to 100 pupil translations, in 100 non-linear steps.
- Minimum baseline δ_x between pupils is set at 4 pixels, or 1/24th of a pupil diameter
- Maximum baseline between pupils m_x. The width of the initial row of pupils is set at 4 pupil translations. The grid of saved pupil positions is a square.
- Outer scale in terms of pupil diameter. This cannot be rescaled for as it is an inherent property of the phase screens used to generate the Zernike coefficients. Instead three copies of the covariance library are saved, with an $L0/D$ ratio of 1, 3 and 10 respectively.

Due to the parameters chosen there is a ±2.4% error in the spatial separations of the covariance matrices and a ±0.8% error in the temporal averaging of the variance, inherent to the covariance library. We have shown that we can extend the time averaging of the grid for all modes. Spatial separations can be accurately extended off the grid for radial orders 2 and above, but not for tip and tilt. The model assumes Taylor's frozen flow and only consider fully developed atmospheric turbulence - dome seeing is not addressed.

To use the generated covariance library we first require the pupil separation between two lines of sight. This is calculated from their angular separation, and the height of the atmospheric layer in question. Next the inverse of the wind direction is applied to the pupil separation. These new coordinates are used to search the library for the two nearest grid points corresponding to this offset. Linear interpolation is used to find the covariance matrix that would exist in between these 4 points. The positive wind direction is then applied as a coordinate transformation. The matrix can then be simply re-scaled for the correct pupil diameter and Fried parameter. This method allows for rapid generation of covariance matrices, even for large telescopes with complex, multi-layer atmospheres.

Chapter 5

Active Optics Correction Using an Off-Axis Guide Star

5.1 Introduction

An active optics system will use a bright guide star as a reference to correct the telescope surface shape (Babcock, 1953). Residual WFE due to the finite exposure time of the AcO WFS is an additional term that is difficult to separate from a telescope misalignment. The model developed in this thesis places a fundamental limit on AcO performance. As discussed in chapter 3 the time required to reduce the atmospheric error from an individual Zernike mode is dependant on both wind direction and telescope diameter. We have tested the effects of finite averaging time in an active optics telescope model. If a single off axis wavefront sensor is used to correct the on-axis line of sight, the covariance library detailed in chapter 4, allowed us to find the potential residual WFE of this correction. We have used this model for both existing telescopes and ELT scales.

Both the Gemini telescopes and the VLT unit telescopes have a primary mirror of 8.1m in diameter, but have very different Field of View (FoV). We have also considered the VISTA telescope. The diameters, FoV and AcO update rates of the three telescopes are detailed in table 5.1. Also included is the outer scale used

for the atmosphere when calculating the residual error for these telescopes. As mentioned in section 4.3.1, we have a full covariance grid for diameter to outer scale ratios of 1, 3 and 10. To be as close as possible to the 39m outer scale for our ELT simulations, we choose a ratio of 3 for our 8m telescopes and 10 for the 4m.

Name	Diameter /m	FoV /degrees	AcO Update Rate /s	Outer Scale Modelled /m
VISTA	4	0.75	40	40
VLT	8.1	1	30	24.3
Gemini	8.1	0.08	30-60	24.3
ELT	39	0.0083	<300	39

Table 5.1: Diameter in metres, FoV in degrees, update rate in seconds and the Outer Scale in metres used in the corresponding simulation for the VISTA, VLT, Gemini and ELT telescopes. ELT will update every 5 minutes in a worst case (Bonnet et al., 2018).

In section 5.2 we explore the residual difference between one time averaged off axis line of sight used to correct instantaneous on-axis line of sight. We discuss our method in section 5.2 and have used a single layer atmospheric model to verify this against the existing ANGuS Monte-Carlo model in section 5.3. Then in 5.4 we have utilised the full ESO 35 layer turbulence profile, as detailed in table 2.1, to explore how variations in the exposure time used in the off-axis line of sight effects the residual difference.

5.2 Theory of Calculating the Residual Variance Error

In this section we detail our theory used to predict the long exposure performance of an AcO system. An off axis WFS measures the atmosphere for an exposure time T. This is then used as a correction for a subsequent on-axis observation. This was an ideal simulation in which to use our covariance library as it required both instantaneous and time averaged covariance matrices for variable pupil separations and wind profiles.

To simulate an AcO system we first measured instantaneous phase along reference line of sight $\phi_{ref}(t)$ for time between T_0 and T_1. This is used to correct the instantaneous phase along the on axis line of sight $\phi_{on}(t)$ between T_1 and T_2. Both processes take an equal amount of time, $T_1 - T_0 = T_2 - T_1$ and ϕ_T represents the telescope error, which is seen by both lines of sight. Both are sampled in n steps. This is shown as a diagram in figure 5.1.

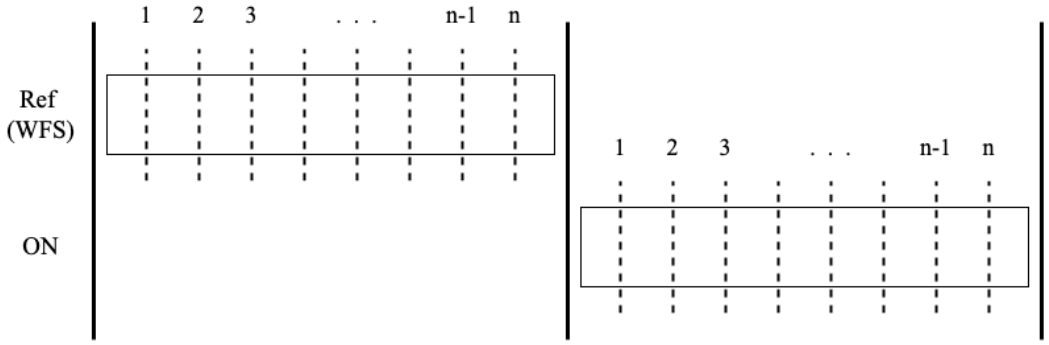


Figure 5.1: Timeline for measuring the reference line of sight between T_0 and T_1 and then correcting the on-axis wavefront between T_1 and T_2

If it is not an instantaneous phase value, and has been averaged over a length of time, we denote this with $\overline{\phi}$ such as:

$$\overline{\phi}_x = \frac{\sum_{k=1}^{n} \phi_x(t_k)}{n}. \tag{5.1}$$

The measured AcO uses the $\overline{\phi}_{ref}$ average between T_0 and T_1. The residual WFE after applying this correction is

$$\phi_{res}(t) = \phi_{on}(t) + \phi_T(t) - \overline{\phi}_{ref} - \overline{\phi}_T. \tag{5.2}$$

In this example we assume that $\phi_T(t) - \overline{\phi}_T \approx 0$ so that the telescope error varies slowly enough it doesnt change over this timeline and therefore cancels out. We are then left with

$$\phi_{res}(t) = \phi_{on}(t) - \overline{\phi}_{ref}. \tag{5.3}$$

The average residual error due to the atmosphere can then be shown as

$$\langle \phi_{res}(t)^2 \rangle = \langle (\phi_{on}(t) - \overline{\phi}_{ref})^2 \rangle, \tag{5.4}$$

which expands out to:

$$\langle \phi_{res}(t)^2 \rangle = \langle \phi_{on}(t)^2 \rangle + \langle \overline{\phi}_{ref}^2 \rangle - 2\langle \phi_{on}(t)\overline{\phi}_{ref} \rangle. \tag{5.5}$$

Now concentrating on that final term,

$$-2\langle \phi_{on}(t)\overline{\phi}_{ref} \rangle = -2\left\langle \frac{\sum_{j=n+1}^{2n} \phi_{on}(t_j)}{n} \frac{\sum_{i=1}^{n} \phi_{ref}(t_i)}{n} \right\rangle. \tag{5.6}$$

As ϕ_{on} and ϕ_{ref} are measured at times t_j and t_i respectively, we use δ to express the time delay as

$$\delta = j - i, \tag{5.7}$$

This allowed for the covariance cross term to be expressed in a way which allowed us to utilise the covariance library. Equation 5.6 expands into n^2 terms which can be grouped together in terms of their δ values:

$$\frac{-2}{n^2} \begin{pmatrix} 1\langle \phi_{on}(t_{i+1})\phi_{ref}(t_i) \rangle & \delta = 1 \\ +2\langle \phi_{on}(t_{i+2})\phi_{ref}(t_i) \rangle & \delta = 2 \\ +3\langle \phi_{on}(t_{i+3})\phi_{ref}(t_i) \rangle & \delta = 3 \\ \cdots & \cdots \\ +n\langle \phi_{on}(t_{i+n})\phi_{ref}(t_i) \rangle & \delta = n \\ +(n-1)\langle \phi_{on}(t_{i+n+1})\phi_{ref}(t_i) \rangle & \delta = n+1 \\ \cdots & \cdots \\ +1\langle \phi_{on}(t_{i+2n-1})\phi_{ref}(t_i) \rangle & \delta = 2n-1 \end{pmatrix} \tag{5.8}$$

This can be expressed as a sum as δ increases from 1 to $2n-1$

$$-2\langle \phi_{on}(t)\overline{\phi}_{ref} \rangle = \frac{-2}{n^2} \sum_{\delta=1}^{2n-1} (n - |\delta - n|)\langle \phi_{on}(t_{i+\delta})\phi_{ref}(t_i) \rangle. \tag{5.9}$$

Substituting equation 5.9 back into 5.5 gives

$$\langle \phi_{res}(t)^2 \rangle = \langle \phi_{on}(t)^2 \rangle + \langle \overline{\phi}_{ref}^2 \rangle - \frac{2}{n^2} \sum_{\delta=1}^{2n-1} (n - |\delta - n|)\langle \phi_{on}(t_{i+\delta})\phi_{ref}(t_i) \rangle. \tag{5.10}$$

This leaves us with three terms to calculate the residual variance after the correction, whilst accounting for the finite exposure of the AcO WFS; the instantaneous variance from the on-axis pupil, a time averaged variance from the off-axis reference pupil and a covariance cross-term, all of which are available in our covariance library.

5.2.1 Method

We utilised the covariance library from Chapter 4 to find the residual variance as calculated in equation 5.10:

- For a given layer of turbulence, this required us to define the angle between the projected pupils, the height and wind velocity.
- Once the angular separation has been converted to a coordinate separation, the rotation transformation for the inverse wind direction is applied to the coordinates.
- The covariance matrix for the new coordinate separation was pulled from the library and the coordinate rotation of the positive wind direction was applied.
- The matrix was then scaled for pupil diameter and Fried parameter as $D/r_0^{5/3}$

The time averaged covariances in the library all represent the same averaging distance in terms of pupil diameter, but this translates to different times in seconds, for different speeds. When utilising a multi-layer profile with variable wind speeds the summed covariance matrices must all represent the same timescales. This limits our covariance model to a total exposure time of 120 seconds at ELT scales.

For each layer in the simulation, we determine the separation of the two pupils based on the angular separation of the two lines of sight. An additional spatial offset is then added to emulate the effect of a time delay between t_i and t_j in frozen flow. After accounting for any rotations due to wind and pupil orientation, we then sum covariance values to simulate long exposures.

5.3 Single Layer Atmosphere and Verification of Theoretical Model

As a check of this theory, we have compared to an independent model, historically named for "All Natural GUide Stars" or ANGuS simulation developed by T.Butterley, previously used in Osborn et al. (2012) and currently used for MOSAIC ground layer adaptive optics simulations. This is an end-to-end Monte Carlo adaptive optics simulation, implemented in Python 3 and C. This comparison utilised ANGuS's multiple layered atmosphere model made of periodic phase screens and its Zernike WFS and Deformable Mirror (DM). The ANGuS model uses a large periodic phase screen generated using the Ellerbroek method, detailed in section 2.1.1, unlike the covariance library which was generated using AOTools infinite phase screens.

Both models ran for a single layer of turbulence with a pupil diameter and outer scale of 4 m, Fried parameter of 0.2 m, wind speed of 2 ms^{-1}, layer height of 50 m and number of steps n=50. ANGuS measures Zernike coefficients along both lines of sight for n steps and applies a correction using the off-axis coefficients to the on-axis pupil for the next n steps and so on. Wavefront measurements from this simulation were measured and produced a series of "uncorrected and corrected" Zernike coefficients, from which we can replicate all terms of equation 5.10.

The variance of the on-axis corrected Zernike coefficients should match our $\langle \phi_{res}(t)^2 \rangle$, providing an initial check on the similarity of the models. The instantaneous variance of the on-axis uncorrected Zernike coefficients should match our first term $\langle \phi_{on}(t)^2 \rangle$, proving the atmosphere we are using is the same. The remaining two terms can be calculated from the results of the ANGuS simulation. The second term can be calculated as the time averaged variance of the Zernike coefficients from the off-axis line of sight. Then finally the cross term can be checked by rearranging

equation 5.10 as:

$$-\frac{2}{n^2}\sum_{\delta=1}^{2n-1}(n-|\delta-n|)\langle\phi_{on}(t_{i+\delta})\phi_{ref}(t_i)\rangle = \langle\phi_{res}(t)^2\rangle - \langle\phi_{on}(t)^2\rangle - \langle\overline{\phi}_{ref}^2\rangle. \quad (5.11)$$

The ANGuS simulation uses finite phase screens so running for a similar number of Zernike coefficients to generate the covariance library was impractical in terms of both storage space and time. Instead it ran each angular variation for 40 independent runs of 3200 coefficients, covering 160 seconds each. This additionally allowed for standard deviation errorbars on the ANGuS coefficients by treating each iteration as a subset of the data as a whole. This subsetting error calculation is similar to the jackknife method, detailed in Efron (1979).

We verify the similarities between these two methods for phase motion both perpendicular and in line with pupil separation respectively in sections 5.3.1 and 5.3.2. For clarity these pupil separations with respect to the phase motion are shown in figure 5.2. These represent the positions of the pupils on the phase screen for the ANGuS simulation, and the separation between the pupils used to select a covariance matrix from our library.

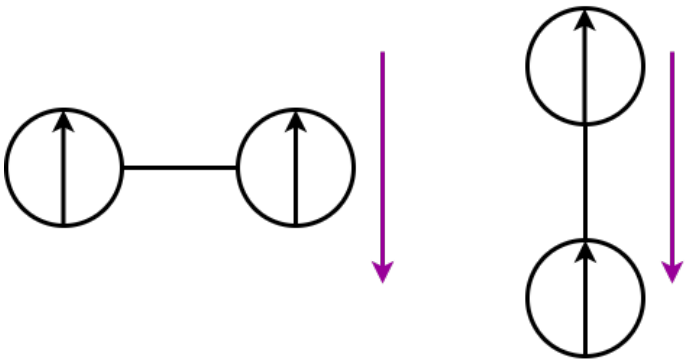


Figure 5.2: Demonstrating the difference in phase motion shown in purple which is perpendicular and in line with the separation between two pupils. Zernike mode orientation is shown in the black lines.

5.3.1 Comparison to ANGuS Model for Separation Perpendicular to Phase Motion

To check the accuracy of the method proposed in equation 5.10 we compared to an output of the ANGuS simulation which should represent the same residual error but produced in a different way. In order for the comparison to work, the pupil separation, Zernike modes and phase motion had to be aligned with each other. For both models the phase moves in the same direction as tip. We tested two different directions of pupil separation: parallel and perpendicular to the phase motion. In order to fully test the accuracy of our simulation we ran these comparisons for a short simulated exposure time of 0.25 seconds and a longer one of 2 seconds.

Figure 5.3 shows the three terms from equation 5.10 as well as the calculated residual variance term for the modes 2 and 4, tip and focus, with a short $T_1 - T_0$ of 0.25 seconds. We chose to examine modes tip and focus here as results in Chapter 3 have shown tip to be noticeably affected by changes to the wind direction, particularly over 90 degrees, which would indicate a misalignment between the two simulations. Focus on the other hand is unaffected by angular changes and should indicate a provide a more general indicator in the accuracy of active optics simulation.

The blue lines showing the instantaneous variance match very well - the two modelled atmospheres are in good agreement. Both variances are unaffected by the change in pupil separation as expected. There is a slight but consistent discrepancy between the time averaged variances which is discussed in section 5.3.3. The covariance cross term fluctuates before settling around zero, in a similar fashion the results from Wilson and Jenkins (1996) and section 4.3.2.1. The cross terms from both models are well matched which suggests that equation 5.10 accurately describes the time averaging and temporal delay. The theoretical model results calculate the three terms individually, whereas the cross term from the ANGuS model is the remainder from the corrected term and the two variances, indicating

that both our derivation and implementation is correct.

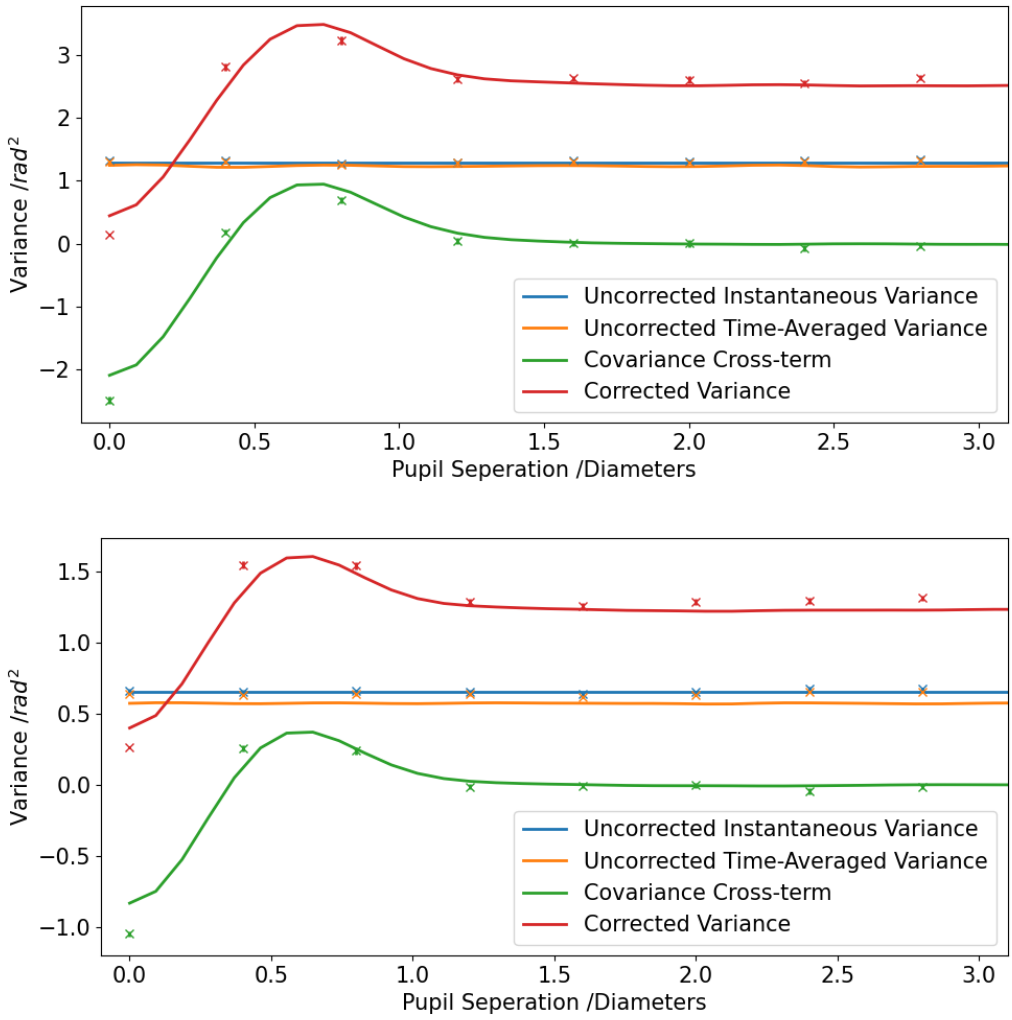


Figure 5.3: Instantaneous $\langle \phi_{on}(t)^2 \rangle$ and time averaged variance $\langle \overline{\phi}^2_{ref} \rangle$, covariance cross term $2\langle \phi_{on}(t)\overline{\phi}_{ref} \rangle$ and corrected variance $\langle \phi_{res}(t)^2 \rangle$ as separation between the pupils increases perpendicular to phase motion. Solid lines are from the theoretical model, the individual points are from the ANGuS simulation. Both simulations ran using a single layer of von Karman turbulence with $D = L_0 = 4$m and r0 of 0.2m moving at $2ms^{-1}$, for a short exposure time of 0.25 seconds. Results are shown for z2 (tip) and z4 (focus) top and bottom respectively.

Figure 5.4 shows the uncorrected, instantaneous variance $\langle \phi_{on}(t)^2 \rangle$ and the corrected variance $\langle \phi_{res}(t)^2 \rangle$ for a longer exposure time of 2 seconds. Values from the corrected variances from our theoretical model are not consistently within the error bounds of the results from the ANGuS simulation, but do have the same scale

and approximate shape. The consistent disparity between the corrected variance results from both models stems from the time averaged variance term. Given the dependence on wind direction for the time averaged Zernike variance established in Chapter 3, we can conclude that the phase is moving in the same direction for both models, relative to the Zernike modes. For this longer exposure time the covariance cross terms still vary between baselines of 0 and $\sim$1 pupils, producing the peak in the corrected variance, but with a much smaller amplitude. This is expected given that the same n steps are representing a longer separation and therefore smaller individual covariances.

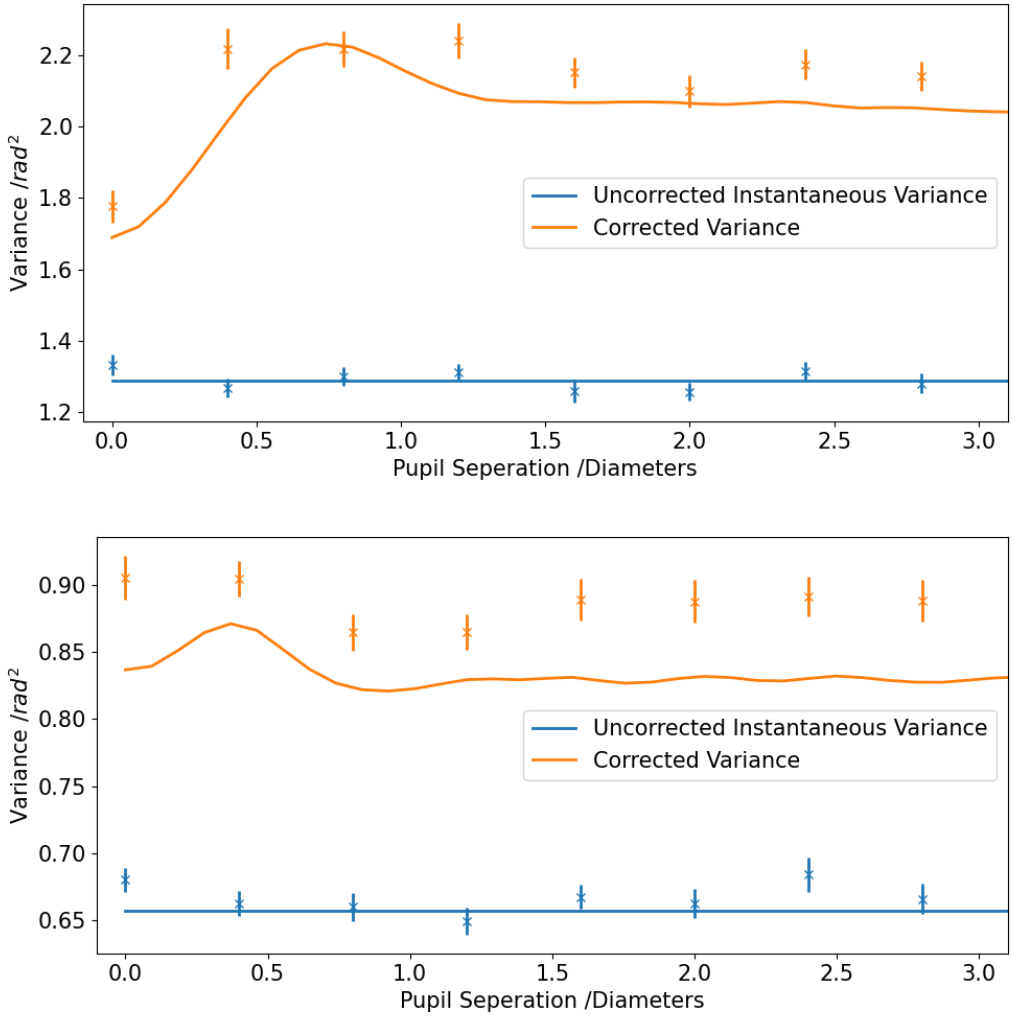


Figure 5.4: Instantaneous $\langle \phi_{on}(t)^2 \rangle$ and corrected variance $\langle \phi_{res}(t)^2 \rangle$ as separation between the pupils increases perpendicular to phase motion. Solid lines are from the theoretical model, the individual points are from the ANGuS simulation. Both simulations ran using a single layer of von Karman turbulence with $D = L_0 = 4$m and r0 of 0.2m moving at $2ms^{-1}$, for a longer exposure time of 2 seconds. Results are shown for Z2 (tip) and Z4 (focus) top and bottom respectively.

5.3.2 Comparison to ANGuS Model for Separation In-Line with Phase Motion

The other variation to consider as part of this verification is the direction of separation with respect to the phase motion. This allows us to further confirm the orientations of the pupil separations and test the effects this has on residual variance in a single layer model.

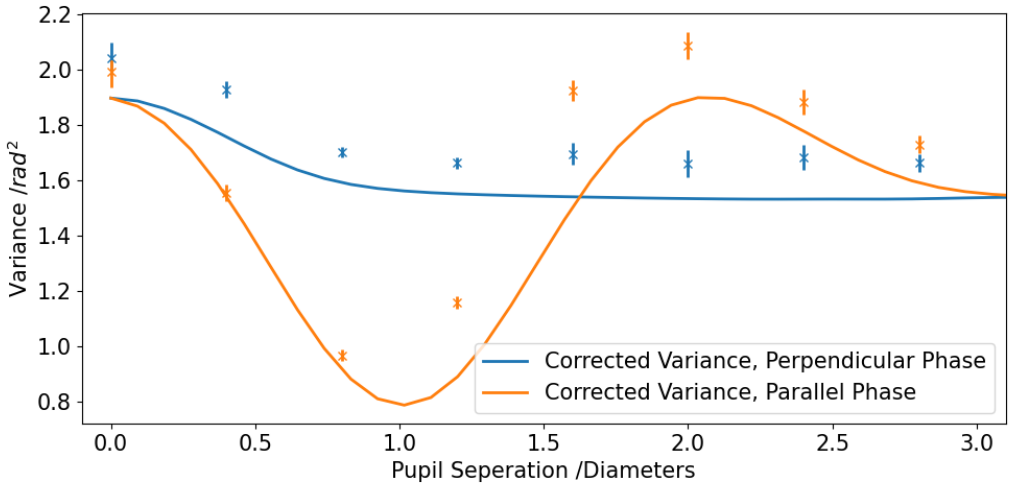


Figure 5.5: Corrected variance $\langle \phi_{res}(t)^2 \rangle$ for Z3, tilt, as separation between the pupils increases both parallel and perpendicular to phase motion. Solid lines are from the theoretical model, the individual points are from the ANGuS simulation. Both simulations ran using a single layer of von Karman turbulence with pupil diameter and outer scale L_0 = 4m and r0 of 0.2m moving at 2 ms^{-1}, for a longer exposure time of 2 seconds.

Starting with mode tilt in figure 5.5, whilst the scales are slightly different we see that the trends/pattern for the corrected variance term of modes tip and tilt are reversed compared to figure 5.3 - it would seem that the covariances are similarly affected by a 90 degree change in separation as the Zernike variances are by a 90 degree phase rotation as discussed in section 3.2.2. There is a small but consistent discrepancy between the residual variance from our theoretical model and the ANGuS simulation. This stems from the time averaged variance term and is discussed in 5.3.3.

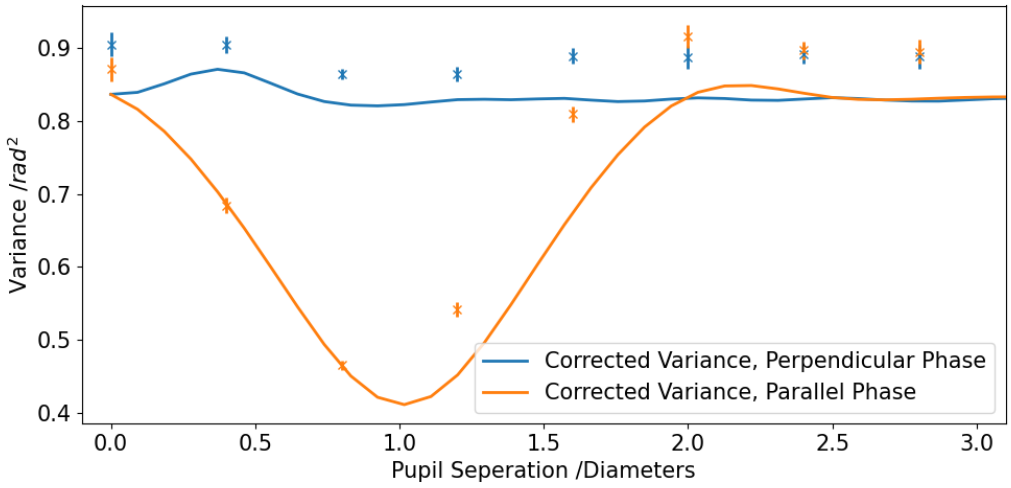


Figure 5.6: Corrected variance $\langle \phi_{res}(t)^2 \rangle$ for z4, focus, as separation between the pupils increases both parallel and perpendicular to phase motion. Solid lines are from the theoretical model, the individual points are from the ANGuS simulation. Both simulations ran using a single layer of von Karman turbulence with pupil diameter and outer scale L_0 = 4m and r0 of 0.2m moving at 2 ms^{-1}, for a longer exposure time of 2 seconds.

Changes to the direction of pupil separation also cause variation in the corrected variance for mode focus. Focus has an azimuthal order of 0 and when considering a singular line of sight was unaffected by changes to the wind direction. Figure 5.6 shows the residual variance for mode focus after an exposure time of 2 seconds. Given our use of Taylor's frozen flow, the off-axis reference line of sight is measuring from phase that will be seen by the on-axis pupil at a later time. This effect is reduced for wider off-axis separations as there is reduced overlap in the instantaneous covariance values within the same time frame, The reduction in overlap means the residual variance for all modes will be affected by relative changes between angular separation of the pupils and wind direction. Beyond a separation of $\sim$ 2 pupil diameters the residual variances for both of these separations converge.

5.3.3 Discrepancies in the Time Averaged Variance

To further explore the differences in the time averaged variance $\langle \overline{\phi}^2_{ref} \rangle$ we can compare both models to the results presented in Gordon et al. (2011). We can re-scale the ANGuS results to match the 30m diameter, $r0 = 0.14m$, $v = 10ms^{-1}$ and compare to Gordon figure 3b, which shows the effects of different outer scales. We have plotted the ±3% error from the covariance library on the results from our theoretical model.

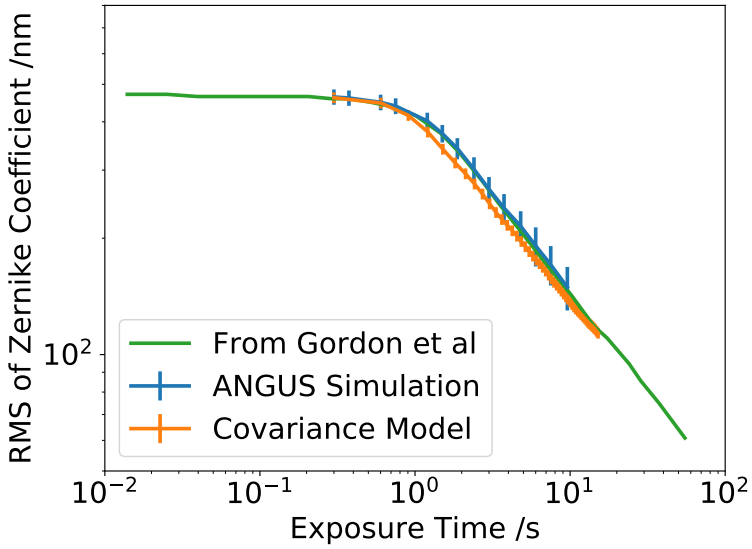


Figure 5.7: Time-averaged RMS WFE for frozen flow atmospheric turbulence for Z4, focus, using a single layer of von Karman turbulence with outer scale $L_0 =$ 39 m and r_0 of 0.14 m moving at 10 ms^{-1} for an ELT sized aperture of diameter 39 m. Comparison between our covariance based simulation, the ANGuS model and the results presented in Gordon et al. (2011)

Figure 5.7 shows that both models match the results from Gordon reasonably well, and at this ELT scale are within error bounds of one another, but that the ANGuS simulation is closer to the results from Gordon et al. (2011). One possible explanation for this is that both Gordon and ANGuS both use finite sized phase screens, whereas our theoretical model uses an infinite phase screen and there may be a systematic difference between the two. There may be a discrepancy in the

outer scale values between the simulations. Alternatively if there is any variation in pixel scales then it is possible that the variances do match well as shown in figure 5.7, we are simply comparing two slightly different specific times in .

5.4 Realistic Atmospheric Model and Telescopes

Having built confidence that our model works as intended, we then started utilising the full extents of our covariance grid, working again with a multi-layer atmospheric profile with a pupil diameter of 30m.

5.4.1 Exploring Effects of Wind Speed and Strength of a Multiple Layered Atmosphere

To clearly show the effects of the different speeds and weights of the different layers we started with a 35 layer profile of the atmosphere, with the same heights as detailed in Table 2.1, but with all layers moving at the same speed and with equal strengths. Whilst this is a non-physical profile, it allowed for the effects of the variable strengths and speeds to be seen more clearly in comparison. A short, AO like exposure time of 0.25 seconds was chosen in this section, to amplify the effects of increasing pupil separation. The theory behind the AcO and AO corrections is same, just at different rates. The AO updates very rapidly and is able to correct for the atmospheric aberrations and AcO updates slowly, to allow for correction of the telescopic errors.

As we are now considering multiple layers we must express the distance between the on and off-axis lines of sight not in pupil diameters but in terms of their angular separation. In this section we increase from 0 degrees (perfectly overlapped) to 0.1 degrees or 6 arcminutes. For mode focus we plot the uncorrected instantaneous variance and the corrected variance for 4 different atmospheric profiles in figure 5.8.

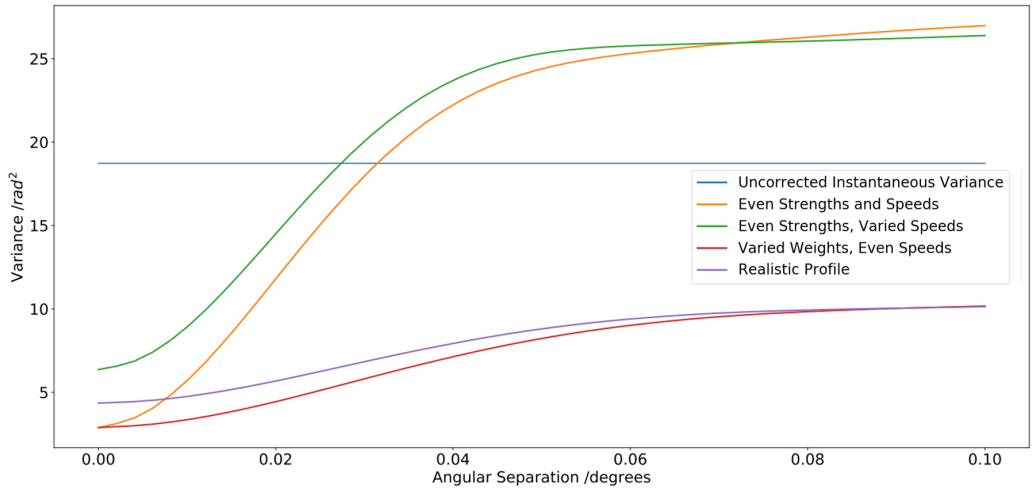


Figure 5.8: The residual focus variance term for multi-layered atmospheric model with angular separation increasing from 0 to 0.1 degrees. This is shown for three simplified atmospheric models with constant speeds and/or strengths across the layers and one with both varying according to the ESO 35 layer profile. All use layers use von Karman turbulence with $L_0 = 39$m and r_0 of 0.157m for an ELT sized aperture of diameter 39m. Exposure time of 0.25s is used, putting this in the adaptive optics regime.

First we consider a basic case, a wind speed of 5.1ms^{-1}, the slowest speed of the profile listed in table 2.1, with all layers given equal strengths - the orange curve in figure 5.8. The Zernike variances for a single line of sight are not affected by position within a phase screen or our covariance library, so as we sum together the $\langle \phi_{on}(t)^2 \rangle$ and $\langle \overline{\phi}^2_{ref} \rangle$ terms across the 35 layers, this should be equivalent to a single layer profile as used in section 5.3. The cross term $2\langle \phi_{on}(t)\overline{\phi}_{ref} \rangle$ encodes the covariance between the two lines of sight and so does change with the new multi-layered profile. Increasing the angular separation leads to a larger physical separation for layers at higher altitude. As we sum over layers of increasing height we effectively sum over increasing physical separation, increasing the anisoplanatism effects. This leads to different covariance cross term results from each atmospheric layer regardless of wind speed so the covariance, and therefore the residual variance, will not be the same as for a single layer model with the same outer scale, Fried parameter and wind speeds. The performance deteriorates rapidly with increasing angular separation due to the top heavy profile, the same as Single-Conjugated Adaptive

Optics (SCAO).

Next we alter the speeds of the different layers to match the ESO 35 layer profile, as shown in table 2.1. This does not effect the instantaneous $\langle \phi_{on}(t)^2 \rangle$, but will effect the time averaged $\langle \overline{\phi}^2_{ref} \rangle$ and the covariance cross term $2\langle \phi_{on}(t)\overline{\phi}_{ref} \rangle$. In this scenario all the layers still have equal strength, so the time averaged variance will be equivalent to a single layer profile with a speed of $11ms^{-1}$, this is the average across the ESO profile. We therefore expect the time averaged variance term to be lowered. Whilst the covariance cross term is made of only instantaneous values it is affected by wind speed as it changes the distance travelled between the instantaneous measurements - faster travelling layers will generate smaller covariance cross terms. We see this in the orange curve of figure 5.8 where the faster model has a slightly higher residual variance for small pupil separations. Overall this has an increased wind speed, hence the worse performance. As the covariance cross term is approaches zero as pupil separation increases the increased speed has a less noticeable effect.

There is a much more pronounced effect when changing the strengths of the layers shown in figure 5.8. As mentioned before, 70% of the overall strength of the profile lies in the lowermost eight layers. This emphasis on the smaller separations increases the magnitude of the overall cross-covariance term regardless of the angular distance, which in turn decreases the residual error. The purple line in figure 5.8 shows the results of a realistic atmospheric model with layers with varied strengths and speeds. Again there is a small increase in residual error for small separations caused by increasing the speeds of most of the layers. The atmospheric AcO error is clearly ground dominated, demonstrating the importance of detailed ground layer Cn^2 and $\vec{v}(h)$ profiling for this application.

5.4.2 Variations in Exposure Time

The main consideration for our active optics model, and the focus of this section, is the exposure time used to measure the off axis line of sight $\langle\overline{\phi}^2_{ref}\rangle$ before correcting; and how this effects the residual on-axis error. This is representational of the fundamental limit on AcO performance with one off-axis WFS, purely from an atmospheric perspective.

As the exposure time increases the time averaged variance term will decrease approximately as a power law - in line with the results discussed in Chapter 3. Equation 5.9 shows the cross term is made of a sum of instantaneous covariance values that cover increasing δ time differences between the on-axis and reference lines of sight. As the exposure time increases, so will the spatial difference between the lines of sight and the covariance cross term will reduce to zero. We have seen that the covariance between two pupils can be negative; as demonstrated in Chapter 4, showing instantaneous covariance values for increasing pupil separation. The cross term drops quite rapidly for short exposure times of a few seconds to negative values before settling around zero. This trough occurs at shorter timescales for more complex Zernike modes, which will correspond to a peak in the residual variance, as shown in figure 5.9.

Figure 5.9 shows residual variance for azimuthal order -1 modes against exposure times between 0.01 and 30 seconds. For very short timescales the residual variance is smaller than the instantaneous variance. In this regime our model theoretically could correct for the atmosphere. Once the residual variance has crossed the instantaneous variance it peaks, this would be the worst timescale to correct for a particular mode. We note that this peak occurs at shorter timescales for higher order modes, but is of smaller value in comparison. Z3 Tip has a maximum residual error at $\sim$7 seconds - increasing the exposure time after this point will reduce the residual error for all modes.

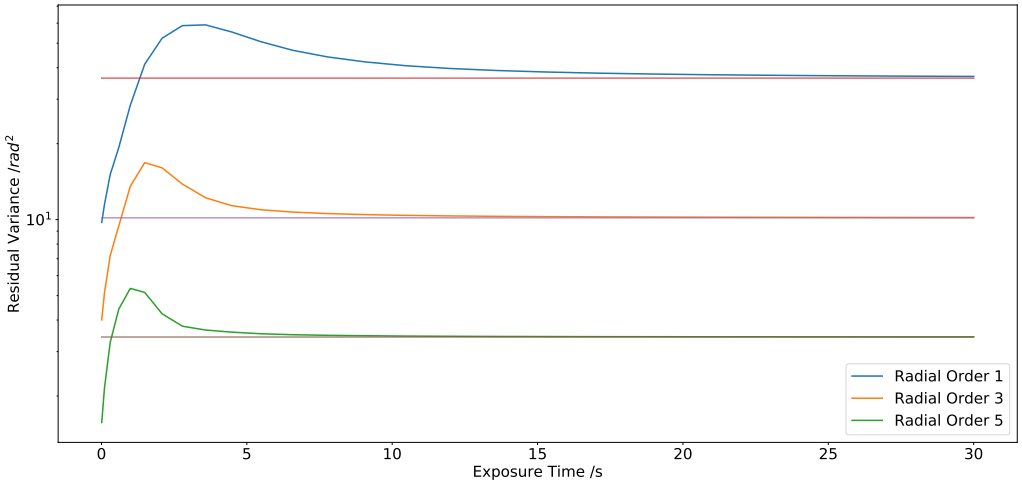


Figure 5.9: The residual $\langle \phi_{res}(t)^2 \rangle$ and insantaneous variance $\langle \phi_{on}(t)^2 \rangle$term for multi-layered atmospheric model with increasing exposure for the off-axis reference pupil. This is shown for three different radial orders with azimuthal order -1. The guide star is 6 arcmin off-axis. Also included is the uncorrected instantaneous variance term for the corresponding mode. Generated using the ESO 35 layer profile with L_0 = 39m and r_0 of 0.157m for an ELT sized aperture of diameter 39m.

5.4.2.1 Variations in Wind Direction

Here we can finally test the impacts of wind velocity on this model. We vary the directions of the wind profile, both helically between the different layers as discussed in Chapter 3, but now also relative to the azimuthal coordinate of the guide star. We have established in section 5.3.2 that the angular difference between phase motion and the pupil separation can have a large effect on the scale of the residual variance for individual modes when considering a single layer model and section 3.3 discusses how, given its relative strength, the motion of the ground layer has a strong effect on the overall averaging times when considering a single line of sight. In this section, we have considered the AcO excess variance $\langle \phi_{exc}(t)^2 \rangle$:

$$\langle \phi_{exc}(t)^2 \rangle = \langle \overline{\phi}_{ref}^2 \rangle - \frac{2}{n^2} \sum_{\delta=1}^{2n-1} (n - |\delta - n|) \langle \phi_{on}(t_{i+\delta}) \phi_{ref}(t_i) \rangle, \tag{5.12}$$

or the difference between the residual error and the instantaneous variance. For simplicity we have varied the ground layer in steps of 90 degrees and shown the

effects on modes with azimuthal order ±1:

- All layers aligned at 0 degrees
- All layers aligned at 90 degrees
- Angular helical dispersion of 360 degrees, starting with the ground layer at 0 degrees
- Angular helical dispersion of 360 degrees, starting with the ground layer at 90 degrees

where 0 degrees is aligned with the slope of tip.

Figures 5.10 and 5.11 show the changes to the AcO excess RMS with exposure time for different angular dispersions. These show the expected patterns in that the results of phase motion at 0 degrees and 90 degrees are reversed for azimuthal orders 1 and -1. Taking the results for Z2, tip, the peak in RMS is reduced for ground layers moving at 0 degrees compared to 90 (∼270 versus 390 nm), but decreases more slowly - for exposure times longer than 8 seconds the residual RMS is on average 100 nm higher for phase motion aligned at 0 degrees. The same logic is true when considering the angular dispersion cases, but the differences are less pronounced. In Chapter 3, when just considering the variance, modes with the same azimuthal order but different radial order decreased with the same approximate power laws, but higher order modes were at smaller scales. Now when considering both variances and covariances, this is no longer completely true. Take mode 7 in figure 5.11 for example, when the wind direction is aligned at 90 degrees, the AcO excess RMS dips and rises again between 2 and 10 seconds, which we do not see in mode 2.

Whilst changes in wind dispersion do affect all individual modes, the effect on the sum total is minimal, as seen in figure 5.12. As seen in Chapter three, pairs of modes average together to produce the same variance, regardless of wind direction. The

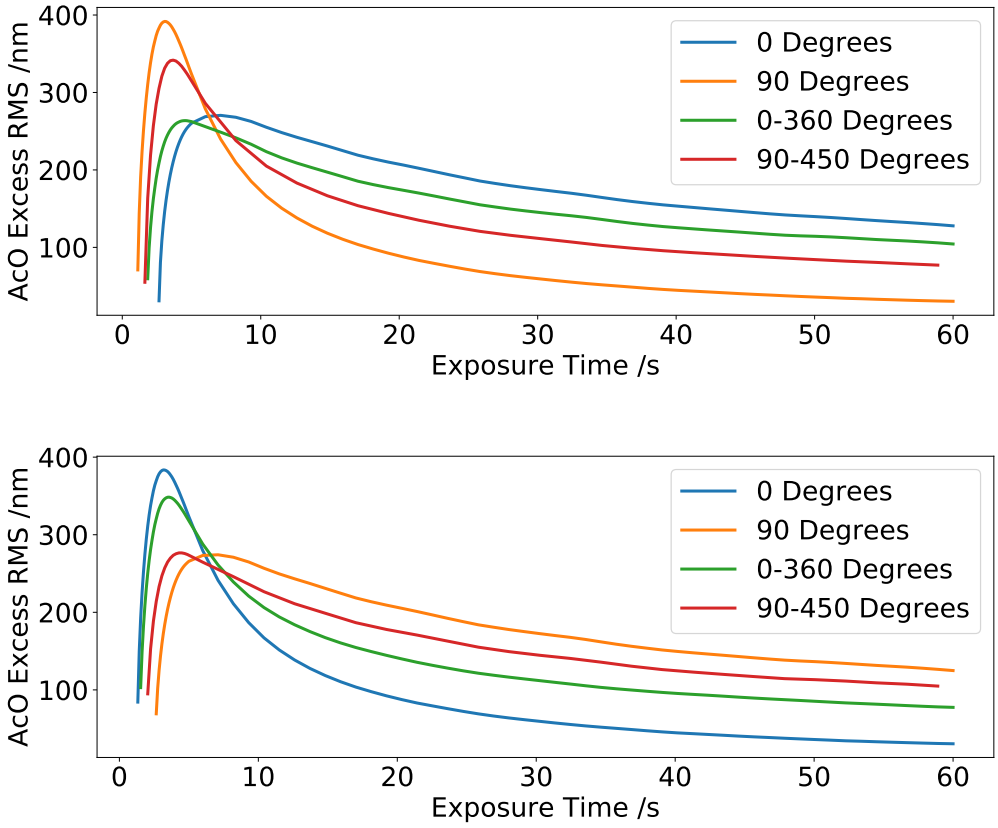


Figure 5.10: The excess AcO RMS term $\langle \phi_{exc}(t) \rangle$ for increasing exposure time, for Zernike modes tip and tilt upper and lower respectively. This is shown for four different velocity dispersions. The guide star is 6 arcmin off-axis. Generated using the ESO 35 layer profile with $L_0 = 39$m and r_0 of 0.157m for an ELT sized aperture of diameter 39m.

same is true of the AcO excess RMS from pairs of modes, so the only variation to the sum of the first ten radial orders stems from the small changes to the azimuthal order 0 modes - of which there are only five. At ELT scales, exposure times should be at least 10 seconds in length, to avoid the peak in AcO excess RMS. To reach 350nm, half the maximum value, takes over 30 seconds. Any shorter will risk extremely large residual errors.

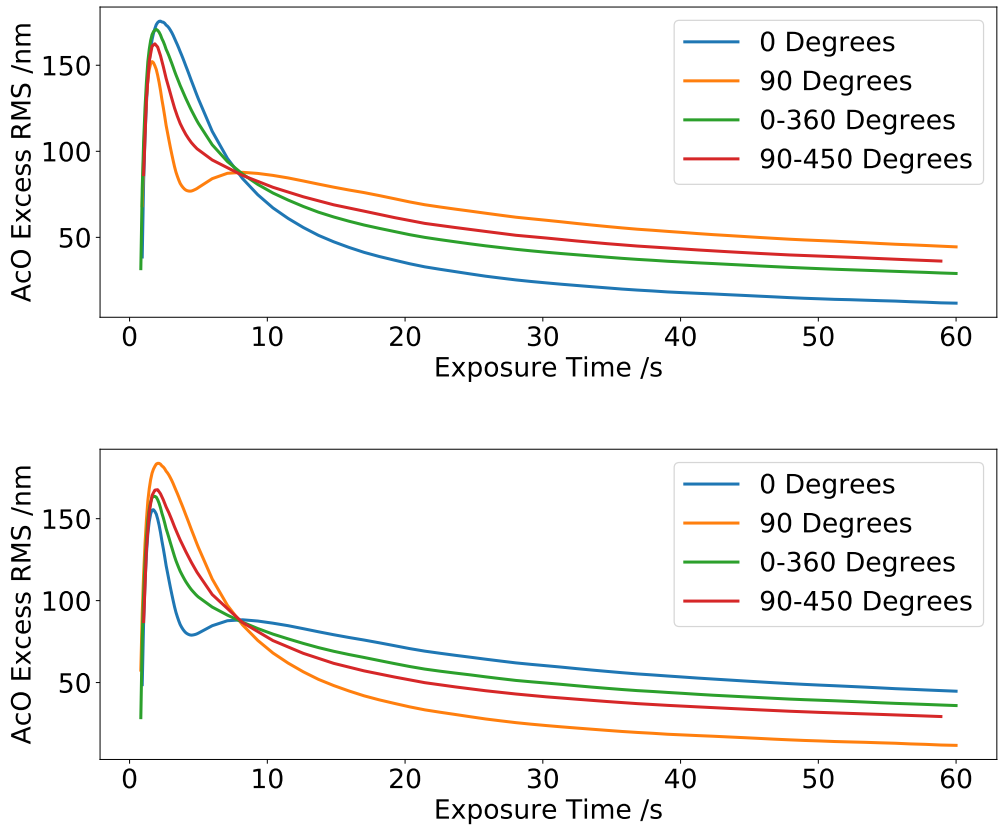


Figure 5.11: The excess AcO RMS term $\langle \phi_{exc}(t) \rangle$ for increasing exposure time, for Zernike modes 7 and 8 upper and lower respectively. This is shown for four different velocity dispersions. The guide star is 6 arcmin off-axis. Generated using the ESO 35 layer profile with $L_0 = 39$m and r_0 of 0.157m for an ELT sized aperture of diameter 39m.

5.4.3 Application to Real Telescope Examples

To provide the ELT-scale results from 5.4.2 with some additional context in this section we have calculated the residual error $\langle \phi_{res}(t)^2 \rangle$ for smaller, existing telescopes - VISTA, Gemini and the VLT. We continue to use the median ESO profile for the atmospheric layers.

Boccas and Vucina (2006) states that Gemini's update rate is variable - in 0.5" seeing the update rate is 30 seconds but will increase to 60 seconds in worse conditions of 1" or above. In figure 5.13 we show the excess AcO RMS for a guide

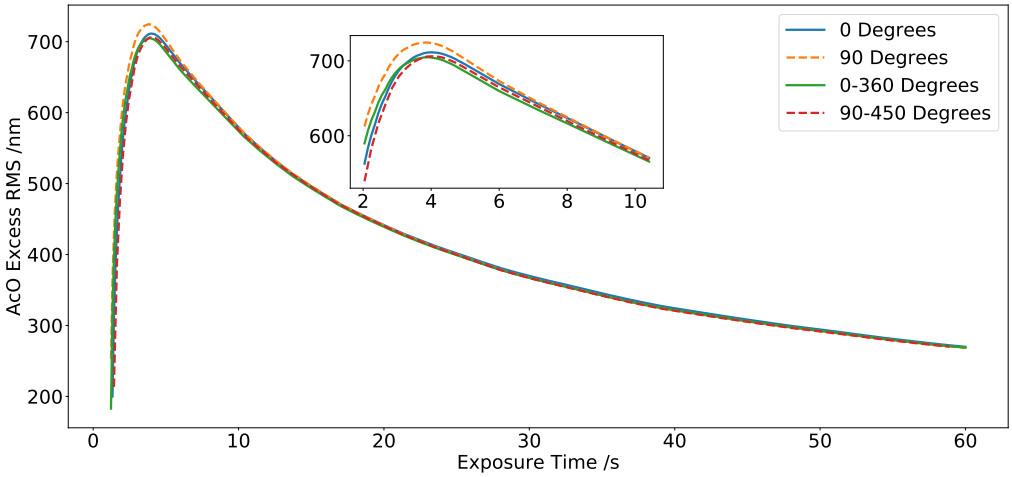


Figure 5.12: The excess AcO RMS term $\langle \phi_{exc}(t) \rangle$ for increasing exposure time, summed together for the different Zernike modes, for increasing exposure time. This is shown for four different velocity dispersions. The guide star is 6 arcmin off-axis. Generated using the ESO 35 layer profile with $L_0 = 39$m and r_0 of 0.157m for an ELT sized aperture of diameter 39m.

star at the edge of the FoV for all three telescopes from table 5.1 and the ELT, for exposure times increasing to 60 seconds. Despite their differences in FoV, the residual covariance for the Gemini and VLT telescopes are only noticeably different for timescales under 10 seconds. The peak residual error is higher for the VLT, which follows the results from 5.4.1 given its wider FoV. However, for longer exposure times the separation in phase is dominated by the temporal split and so both telescopes have a similar error. In the long exposure regime of many seconds, the residual WFE is a function of aperture diameter but not off-axis angle. We also note that the excess AcO RMS for the Gemini telescope only drops from 129 to 98 nm as exposure time increases from 30 to 60 seconds. The residual covariance from VISTA does not peak as highly as the larger telescopes and conversely the ELT peaks at the highest value of 698nm. After an exposure of 60 seconds the excess AcO RMS lowers to 189 nm - equivalent to a 12 second exposure for a Gemini-like pupil separation. In 300 seconds, the "worst case" update rate for the ELT, Bonnet et al. (2018), our model predicts a 90 nm error - just less than the excess AcO RMS for the Gemini 8.1m telescope after a one minutes exposure.

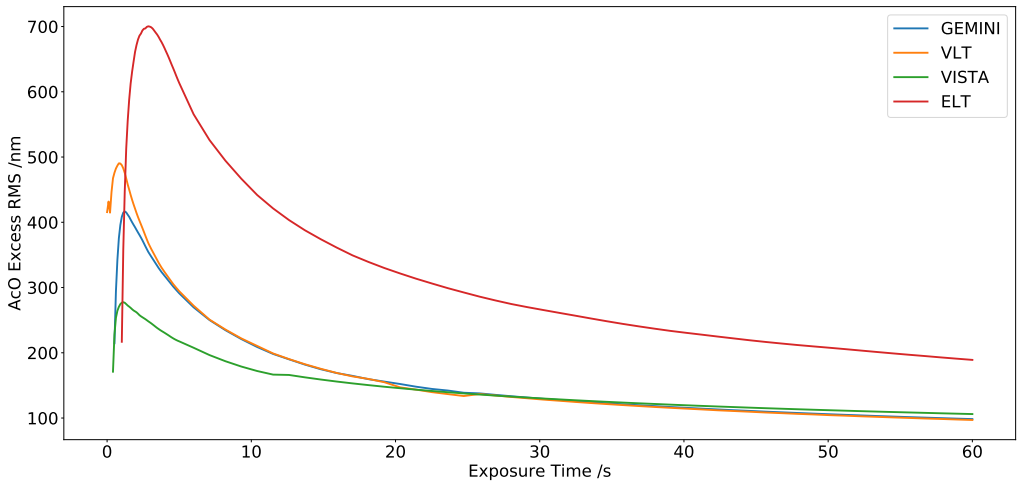


Figure 5.13: The excess AcO RMS sum of the first 10 radial orders with increasing exposure for the off-axis reference pupil. This is shown for Gemini, VLT andVISTA at separations of 0.08, 1 and 0.75 degrees respectively. Generated using the ESO 35 layer profile with r_0 of 0.157m.

In order to balance the need for a frequent update rates (hence shorter exposure times) with a smaller residual error, different modes could, in principle, be averaged for different amounts of time. For example, leaving radial orders one and two at an exposure time of 60 seconds, radial orders three through ten can be shortened to 15 seconds without raising the total excess AcO RMS to above 250 nm. Conversely, if the drive was to shorten the exposure times for the low order modes, radial orders one and two can be measured for 30 seconds and the higher order modes for a minute and achieve the same excess error. If this variable averaging time was viable, the overall wind dispersion becomes an important consideration again. For example taking our 90-450 degree dispersion radial orders one and two measured for 30 seconds and the higher radial orders for a minute gives a total excess AcO RMS of 122 nm. Implementing this would require knowledge of the temporal power spectrum and the variability of the wind velocity profile - and would increase the computational complexity of updating the AcO system. Plantet et al. (2022) presents a method for considering both the spatial and temporal statistics of the turbulence and how this may affect an AO system. This takes the temporal and

spatial covariances of the Zernike modes and produces the temporal cross power spectral densities (Whiteley et al., 1998). They show that only considering spatial covariances causes an SCAO system to overestimate the contributions of anisoplanatism. Implementing a similar analysis using our covariance library would help indicate the viability of a variable averaging time, and is a consideration for future work.

5.5 Conclusions

In this chapter we have used our covariance library to create an AcO model that finds the residual error between one off-axis guide star and the main on-axis line of sight. We have detailed this method in section 5.2, indicating how we have taken advantage of our covariance library to generate a covariance cross term that includes both instantaneous and time averaged values. This theory was verified in sections 5.3.1 and 5.3.2, comparing against an existing Monte Carlo model that was run independently of our theoretical model. Our simulation assumes Taylor's frozen flow, and that the telescope is static.

Confident in the validity of our model we have shown how the residual error peaks at short exposure times before decreasing towards the instantaneous Zernike variance value. We have shown this residual error for a range of telescope sizes. For example at an ELT scale, for a 30 second exposure an off axis wavefront sensor 0.0083 degrees off axis correcting the on axis line of sight will have a excess AcO RMS of 274 nm. For exposure times of 1 to 60 seconds, a single off-axis WFS leads to a relative residual error of on average three times higher error for a 39m telescope vs an 8m diameter. We recommend that exposure times be above 10 seconds at ELT scales.

Future work for this model would again involve varying the wind directions as a function of time to create a more accurate atmospheric profile. In section 5.4.3 we discussed the possibility of varying the exposure times of different modes to achieve the same level of residual error. A useful extension of this would be to calculate

this algorithmically - if we could input a wind velocity profile, which modes we need to keep with a short exposure time and an error threshold, and output a series of necessary exposure times. We could also extend the simulation to consider multiple guide stars, creating a more accurate ELT model.

An alternative avenue would be to follow on the work presented in Holzlöhner et al. (2014). They have presented an alternative AcO control method for the VST. As a survey telescope with a wide FoV, they propose taking advantage of the thousands of stars in the science image and track the telescopic error with the elongation of the stellar PSFs across the field. The pattern of both sizes and ellipticities of the PSFs is compared against an telescopic-aberration-free version in an analytical model. Using 5th-order geometrical optics this model calculates the mirror misalignments and distortions up to nine degrees of freedom. As this uses aberration theory it claims to be faster than ray tracing methods, but it's function requires the atmospheric effects to be fully averaged. They assume an exposure time of one to a few minutes will average out the atmospheric effects - which we could test with our atmospheric active optics model.

Chapter 6

Conclusions

This thesis has considered the long term phase averaging required for ELT scale active optics and how this is affected by wind velocity dispersion. In this chapter we summarise the results presented in this thesis and suggest probable avenues for future research.

6.1 Effects of Wind Velocity Profiles on Turbulence-Induced Quasi-static Aberrations

Chapter 3 focused on the residual phase variance as a function of averaging time for a single ELT scale line of sight.

- The decay rates for individual Zernike modes can be fitted approximately as a power law (after a break point $\tau_{BP} \approx D/v$).
- The gradient of this fit, and therefore the time taken to reach an acceptable threshold of error, is dependent on the azimuthal order of the Zernike mode and the wind direction. For example, for an ELT scale telescope ($r_0 = 0.14$m,$L_0 = 39$m and a wind speed of 10ms^{-1}) the times taken for Zernike modes to reach a residual error of $\lambda/20$ are 10 and 64 seconds in the K-band and 40 and 1016 seconds in the V-band respectively.

- Modes with azimuthal order zero behave in the same way for any wind direction.
- Azimuthal pairs $Z_n \pm m$, two modes within a radial order with $\pm$ azimuthal order m, will switch between power law gradients ~ -0.45 and -1 every $\frac{180}{|m|}$ degrees of wind direction rotation - for a single layer model.
- When considering an atmospheric profile with multiple layers, the effects of wind dispersion are not as simple to quantify given the different strengths and speeds of the individual layers. The averaging times do tend back to the azimuthally averaged case for larger angular dispersions, as this is equivalently averaging over wind direction.
- Given its relative strength, the velocity of the ground layer has a large impact on averaging times in comparison to layers at higher altitudes. Slower moving turbulence exacerbates the split in averaging times between pairs of modes.
- Varying the ground layer wind speeds from 18 to 1 m/s can increase averaging time by a factor of 50.

6.2 Generating A Covariance Library for Two Lines of Sight

In Chapter 4 we have verified the use of a coordinate transform of the Zernike coefficient covariance matrix to rotate the wind velocity. The resultant covariance matrix covariance library allows for more efficient simulation varying wind dispersions for multi-layer atmospheric profiles. The chosen parameter spaces for this library are:

- Maximum baseline between pupils (m_x). The total grid size is set at 4x4 pupil diameters. After this separation the normalised instantaneous covariance is

0 for radial order 2 and above as there is no longer a correlation between the covariance of the two pupils.

- Minimum baseline between pupils (δ_x). This sampling was tested against the normalised instantaneous covariance. δ_x was set at 4 pixels, or 1/24th of a pupil diameter. This sampling is fine enough to successfully interpolate between grid points and generates a worst case $\pm 2.4\%$ error in the instantaneous covariance values.

- Number of Zernike coefficient values to be averaged over T_{av}, up to 100 pupil translations which covered a 100 second exposure at ELT scales for the ESO atmospheric profile. We increased T_{av} in 100 non-linear steps because it limited the storage necessary for the covariance library. This generated a worst case $\pm 0.8\%$ error in the time averaged variance values.

- Number of Zernike modes, fixed at 66 modes, the first 10 radial orders to ensure there was no additional error in our CRCM method for higher order modes.

- By setting $D = r_0$, the covariance matrices can be easily re-scaled to any r_0 value when utilising the covariance library. The covariance matrices can used for variable wind speeds, directions and turbulence strength.

- The outer scale L_0 cannot be rescaled for, as it is an inherent property of the phase screens used to generate the Zernike coefficients. The covariance libraries must be regenerated for specific outer scales. Three copies of the covariance library are saved, with an L_0/D ratio of 1, 3 and 10 respectively.

- The use of this library is only valid for atmospheres exhibiting Taylor's frozen flow (Taylor, 1938). Non-frozen flow turbulence is seen at astronomical sites, but further characterisation of dome seeing is required before fully understanding the accuracy of our assumption and how we might adjust our model to account for this.

6.3 Active Optics Correction Using an Off-Axis Guide Star

Chapter 5 utilised our covariance library as an active optics simulation, to calculate the residual difference after a finite exposure time.

- We have verified our method to calculate the residual error against an independent Monte Carlo simulation (ANGuS).
- For a single layer model, all modes are effected by changes in wind direction as, given our assumption of Taylor's frozen flow, the winds direction may either increase or decrease the separation between the pupils as time passes.
- Increasing the exposure time for our multi-layer atmospheric profile causes the residual difference between two lines of sight to first peak at a short exposure, giving the worst timescale to apply a correction, before slowly decreasing again. As the exposure time increases we move from an AO regime into AcO.
- For long exposures the residual WFE is a function of aperture diameter and C_n^2, but not the off-axis angle.
- The exposure time to reach an acceptable level of residual error for an ELT scale telescope will be in the hundreds of seconds rather than the tens required for existing telescopes.
- The wind dispersion has a less noticeable effect when considering the sum error over multiple radial orders. It may still be a useful consideration if using a variable averaging time for different modes.

6.4 Future Work

The continued use of the covariance library holds considerable possibilities. The use of real velocity data as a function of time would allow us to evolve the atmospheric

wind velocity dispersion as exposure time increases. An additional possibility would be to run additional copies of the covariance library with different outer scales. Given our focus on ELT scales, $L_0 < D$ would also be a useful inclusion for more accurate modelling of the ground layer.

In simulation we can test how different wind velocity profiles affect the residual error, and how varying the averaging time for different modes might be beneficial for ELT scale telescopes. This could assist not only in decreasing the sum excess AcO RMS at long exposure times, but potentially actively avoid the 10s peak. This may have not been considered before as existing 8m telescopes reach an acceptable level of error within their AcO update rates. At ELT scales the necessary averaging times reach into the minutes, over which time the wind velocity profile may have changed. With ground layer velocity information as a function of time we could more seriously consider the serviceability of varying the modal averaging times.

In Chapter 5 we have considered the residual error between one off-axis line of sight and the main on-axis in an AcO simulation. As telescopes get larger it is more common to have more than one off-axis WFS. The physical separations between the guide stars will be of similar scale to the separation between a guide star and the on-axis target and so our covariance library could also be used to generate wide field long exposure wavefronts (after AcO correction) for multiple lines of sight.

We would approach calculating this residual difference $\langle \overline{\phi}^2_{diff} \rangle$ much the same way as our single line of sight active optics model, but now both lines of sight are time averaged. We would allow for consecutive measurements of different guide stars in our simulation with a continued assumption of Taylor's frozen flow, and add a temporal offset to the second reference line of sight of the appropriate distance given the exposure time and the speed of each atmospheric layer.

The work presented in this thesis only accounts for the atmospheric turbulence error. To progress from this we could test how this will impact ELT performance.

For example examining how an AcO system could account for these atmospheric long term exposure errors. We could investigate how these errors propagate through a tomographic reconstructor simulation - such as ANGuS (previously used in 5.3) - and how the residual difference between the off-axis guide star lines of sight affect the on-axis results.

Using our covariance library we would create a larger covariance matrix that describes multiple lines of sight, each representing an Natural Guide Star (NGS) WFS. This matrix holds both the statistics for the atmosphere used and the separations of the WFS. We would add our residual errors to the ANGuS model via additional Zernike DMs. Our quasistatic errors would act as static aberrations in the reconstructor. This requires Zernike coefficients, so to test our covariance based residual error we would have to convert back to Zernike coefficients via matrix decomposition. Note that as these new coefficients would not be correlated or continuous this method could not be used to make a time sequence of coefficients, but as our initial library includes both instantaneous and time averaged covariance matrices we could recreate Zernike coefficients for different exposure times.

As a multi-active mirror telescope, excess AcO errors may be looped back into the control system of the ELT which may cause focal plane distortions and plate scale errors (Rodeghiero et al., 2021). The results of our simulation could be used as an atmospheric error input into an ELT optical model to test how this could affect ELT performance.

Bibliography

F. Assémat, R. W. Wilson, and E. Gendron. Method for simulating infinitely long and non-stationary phasescreens with optimized memory storage. *Optics Express*, 14(3):988–999, 2006.

M. Azouit and J. Vernin. Optical turbulence profiling with balloons relevant to astronomy and atmospheric physics. *Publications of the Astronomical Society of the Pacific*, 117(831):536, 2005.

H. W. Babcock. The possibility of compensating astronomical seeing. *Publications of the Astronomical Society of the Pacific*, 65(386):229–236, 1953.

M. Boccas and T. Vucina. Performance and upgrades of active optics on gemini telescopes. In *Optomechanical Technologies for Astronomy*, volume 6273, pages 1015–1026. SPIE, 2006.

H. Bonnet, F. Biancat-Marchet, M. Dimmler, M. Esselborn, N. Kornweibel, M. Le Louarn, P.-Y. Madec, E. Marchetti, M. Müller, S. Oberti, et al. Adaptive optics at the eso elt. In *Adaptive Optics Systems VI*, volume 10703, pages 327–335. SPIE, 2018.

J. L. Bufton, P. O. Minott, M. W. Fitzmaurice, and P. J. Titterton. Measurements of turbulence profiles in the troposphere. *JOSA*, 62(9):1068–1070, 1972.

T. Butterley, R. W. Wilson, and M. Sarazin. Determination of the profile of atmospheric optical turbulence strength from slodar data. *Monthly Notices of the Royal Astronomical Society*, 369(2):835–845, 2006.

R. C. Cannon. Optimal bases for wave-front simulation and reconstruction on annular apertures. *JOSA A*, 13(4):862–867, 1996.

Clénet, Y., Conan, J.-M., Fusco, T., and Rousset, G. Preface, 2010. URL `https://doi.org/10.1051/ao4elt/201000001`.

B. Efron. The 1977 rietz lecture. *The annals of Statistics*, 7(1):1–26, 1979.

B. L. Ellerbroek and G. Cochran. Wave optics propagation code for multiconjugate adaptive optics. In *Adaptive Optics Systems and Technology II*, volume 4494, pages 104–120. International Society for Optics and Photonics, 2002.

D. Énard, A. Maréchal, and J. Espiard. Progress in ground-based optical telescopes. *Reports on Progress in Physics*, 59(5):601, 1996.

ESO. ESO Paranal Ambient Conditions Database, 1998. URL `http://archive.eso.org/wdb/wdb/asm/meteo_paranal/form`. Available at `http://archive.eso.org/wdb/wdb/asm/meteo_paranal/form`.

D. L. Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *JOSA*, 56(10):1372–1379, 1966.

D. L. Fried. Anisoplanatism in adaptive optics. *JOSA*, 72(1):52–61, 1982.

D. L. Fried and T. Clark. Extruding Kolmogorov-type phasescreen ribbons. *JOSA*, 25(2):463–468, 2008.

P. Gill, T. Butterley, and T. Morris. Effects of wind velocity profiles on turbulence-induced quasi-static aberrations. In *Adaptive Optics Systems VIII*, volume 12185, pages 1812–1819. SPIE, 2022.

J. A. Gordon, D. F. Buscher, and F. Baron. Long-exposure filtering of turbulence-degraded wavefronts. *Applied Optics*, 50(27):5303–5309, 2011.

S. Guisard, L. Noethe, and J. Spyromilio. Performance of active optics at the vlt. In *Optical Design, Materials, Fabrication, and Maintenance*, volume 4003, pages 154–164. SPIE, 2000.

E. Hecht. *Optics*. Pearson Education India, 2012.

R. Holzlöhner, A. Rakich, L. Noethe, K. Kuijken, and P. Schipani. Fast active optics control of wide-field telescopes based on science image analysis. In *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation*, volume 9151, pages 859–873. SPIE, 2014.

A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 434(1890):9–13, 1991.

V. Lakshminarayanan and A. Fleck. Zernike polynomials: A guide. *Journal of Modern Optics*, 58(7):545–561, 2011.

G. Lombardi, V. Zitelli, S. Ortolani, and M. Pedani. El roque de los muchachos site characteristics. ii. analysis of wind, relative humidity, and air pressure. *Publications of the Astronomical Society of the Pacific*, 119(853):292, 2007.

Marchetti. Relevant Atmospheric Parameters for E-ELT AO Analysis and Simulations, 2015. Accessed: 2022-06-26.

S. Martin, J. Bange, and F. Beyrich. Meteorological profiling of the lower troposphere using the research uav" m 2 av carolo". *Atmospheric Measurement Techniques*, 4(4):705–716, 2011.

I. Newton. A letter of mr. isaac newton, professor of the mathematicks in the university of cambridge; containing his new theory about light and colors: sent by the author to the publisher from cambridge, febr. 6. 1671/72; in order to be communicated to the r. society. *Philosophical transactions of the Royal Society of London*, 6(80):3075–3087, 2014.

L. Noethe. Active optics in modern large optical telescopes. *Progress in optics*, 43 (3), 2002.

R. J. Noll. Zernike polynomials and atmospheric turbulence. *JOSA*, 66(3):207–211, 1976.

J. Osborn and M. Sarazin. Atmospheric turbulence forecasting with a general circulation model for cerro paranal. *Monthly Notices of the Royal Astronomical Society*, 480(1):1278–1299, 2018.

J. Osborn, F. J. D. C. Juez, D. Guzman, T. Butterley, R. Myers, A. Guesalaga, and J. Laine. Using artificial neural networks for open-loop tomography. *Optics express*, 20(3):2420–2434, 2012.

J. Osborn, T. Butterley, M. Townson, A. Reeves, T. Morris, and R. Wilson. Turbulence velocity profiling for high sensitivity and vertical-resolution atmospheric characterisation with stereo-scidar. *Monthly Notices of the Royal Astronomical Society*, page stw2685, 2016.

C. Plantet, G. Carlà, G. Agapito, and L. Busoni. Spatiotemporal statistics of the turbulent piston-removed phase and zernike coefficients for two distinct beams. *JOSA A*, 39(1):17–27, 2022.

B. C. Platt and R. Shack. History and principles of shack-hartmann wavefront sensing, 2001.

C. Rao, W. Jiang, and N. Ling. Spatial and temporal characterization of phase fluctuations in non-Kolmogorov atmospheric turbulence. *Journal of Modern Optics*, 47(6):1111–1126, 2000.

F. Roddier. The effects of atmospheric turbulence in optical astronomy. In *Progress in Optics*, volume 19, pages 281–376. Elsevier, 1981.

F. Roddier. Adaptive optics in astronomy. 1999.

G. Rodeghiero, C. Arcidiacono, J.-U. Pott, S. Perera, G. Pariani, D. Magrin, H. Riechert, M. Glück, E. Gendron, D. Massari, et al. Performance and limitations of using elt and mcao for 50 μas astrometry. *Journal of Astronomical Telescopes, Instruments, and Systems*, 7(3):035005–035005, 2021.

G. Rousset, T. Fusco, F. Assemat, E. Gendron, T. Morris, C. Robert, R. Myers, M. Cohen, N. Dipper, C. Evans, et al. Eagle moao system conceptual design and related technologies. In *Adaptive Optics Systems II*, volume 7736, pages 297–307. SPIE, 2010.

M. Sarazin, M. Le Louarn, J. Ascenso, G. Lombardi, and J. Navarrete. Defining reference turbulence profiles for e-elt ao performance simulations. In *Proceedings of the Third AO4ELT Conference*, volume 13383, 2013.

H. Shepherd, J. Osborn, R. Wilson, T. Butterley, R. Avila, V. Dhillon, and T. Morris. Stereo-scidar: optical turbulence profiling with high sensitivity using a modified scidar instrument. *Monthly Notices of the Royal Astronomical Society*, 437 (4):3568–3577, 2014.

R. Storvold, H. Eide, P. Utley, K. Stamnes, G. Adalsgeirsdottir, D. Lubin, B. Zak, P. Lawson, T. Svendby, R. Moritz, et al. Boundary-layer structure obtained with a tethered balloon system and large-scale observations of the arctic basin obtained with a satellite data acquisition system at the sheba ice camp. In *8th ARM Science Team Meeting, Arizona*, 1998.

W. Swantner and W. W. Chow. Gram-Schmidt orthonormalization of Zernike polynomials for general aperture shapes. *Applied Optics*, 33(10):1832–1837, 1994.

G. I. Taylor. The spectrum of turbulence. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 164(919):476–490, 1938.

D. L. Terrett, N. Bissonauth, V. Graffagnino, M. Stewart, and W. J. Sutherland. Active optics and auto-guiding control for vista. In *Advanced Software, Control,*

and Communication Systems for Astronomy, volume 5496, pages 129–137. SPIE, 2004.

A. Tokovinin, S. Baumont, and J. Vasquez. Statistics of turbulence profile at cerro tololo. *Monthly Notices of the Royal Astronomical Society*, 340(1):52–58, 2003.

M. Townson, O. Farley, G. O. de Xivry, J. Osborn, and A. Reeves. Aotools: a python package for adaptive optics modelling and analysis. *Optics Express*, 27 (22):31316–31329, 2019.

R. K. Tyson. *Introduction to adaptive optics*, volume 41. SPIE press, 2000.

G. C. Valley. Long-and short-term strehl ratios for turbulence with finite inner and outer scales. *Applied optics*, 18(7):984–987, 1979.

J. Vernin and F. Roddier. Experimental determination of two-dimensional spatiotemporal power spectra of stellar light scintillation evidence for a multilayer structure of the air turbulence in the upper troposphere. *JOSA*, 63(3):270–273, 1973.

T. Von Karman. Progress in the statistical theory of turbulence. *Proceedings of the National Academy of Sciences of the United States of America*, 34(11):530, 1948.

M. R. Whiteley, M. C. Roggemann, and B. M. Welsh. Temporal properties of the zernike expansion coefficients of turbulence-induced phase aberrations for aperture and source motion. *JOSA A*, 15(4):993–1005, 1998.

R. Wilson and C. Jenkins. Adaptive optics for astronomy: theoretical performance and limitations. *Monthly Notices of the Royal Astronomical Society*, 278(1):39–61, 1996.

R. Wilson, F. Franza, L. Noethe, and G. Andreoni. Active optics: Iv. set-up and performance of the optics of the eso new technology telescope (ntt) in the observatory. *Journal of modern optics*, 38(2):219–243, 1991.

R. N. Wilson, F. Franza, and L. Noethe. Active optics: I. a system for optimizing the optical quality and reducing the costs of large telescopes. *Journal of modern optics*, 34(4):485–509, 1987.

R. W. Wilson. Slodar: measuring optical turbulence altitude with a shack–hartmann wavefront sensor. *Monthly Notices of the Royal Astronomical Society*, 337(1):103–108, 2002.

D. Winker. Effect of a finite outer scale on the Zernike decomposition of atmospheric optical turbulence. *JOSA*, 8(10):1568–1573, 1991.

Colophon

This thesis is based on a template developed by Matthew Townson and Andrew Reeves. It was typeset with LaTeX 2_ε. It was created using the *memoir* package, maintained by Lars Madsen, with the *madsen* chapter style. The font used is Latin Modern, derived from fonts designed by Donald E. Kunith.