

Dynamic Tomographic Algorithms for Multi-Object Adaptive Optics: Increasing  
sky-coverage by increasing the limiting magnitude for Raven, a science and  
technology demonstrator

by

Kate Jackson

B.Sc., University of British Columbia, 2008

M.A.Sc., University of Victoria, 2010

A Dissertation Submitted in Partial Fulfillment of the  
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mechanical Engineering

© Kate Jackson 2014  
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by  
photocopying or other means, without the permission of the author.

Dynamic Tomographic Algorithms for Multi-Object Adaptive Optics: Increasing  
sky-coverage by increasing the limiting magnitude for Raven, a science and  
technology demonstrator

by

Kate Jackson

B.Sc., University of British Columbia, 2008

M.A.Sc., University of Victoria, 2010

Supervisory Committee

---

Dr. C. Bradley, Supervisor  
(Department of Mechanical Engineering)

---

Dr. C. Correia, Departmental Member  
(Department of Mechanical Engineering)

---

Dr. W.S. Lu, Outside Member  
(Department of Electrical Engineering)

## Supervisory Committee

---

Dr. C. Bradley, Supervisor  
(Department of Mechanical Engineering)

---

Dr. C. Correia, Departmental Member  
(Department of Mechanical Engineering)

---

Dr. W.S. Lu, Outside Member  
(Department of Electrical Engineering)

## ABSTRACT

This dissertation outlines the development of static and dynamic tomographic wave-front (WF) reconstructors tailored to Multi-Object Adaptive Optics (MOAO). They are applied to Raven, the first MOAO science and technology demonstrator recently installed on an 8m telescope, with the goal of increasing the limiting magnitude in order to increase sky coverage. The results of a new minimum mean-square error (MMSE) solution based on spatio-angular (SA) correlation functions are shown, which adopts a zonal representation of the wave-front and its associated signals. This solution is outlined for the static reconstructor and then extended for the use of stand-alone temporal prediction. Furthermore, it is implemented as the prediction model in a pupil plane based Linear Quadratic Gaussian (LQG) algorithm. The algorithms have been fully tested in the laboratory and compared to the results from Monte-Carlo simulations of the Raven system. The simulations indicate that an increase in limiting magnitude of up to one magnitude can be expected when prediction is implemented. Two or more magnitudes of improvement may be achievable when the LQG is used. These results are confirmed by laboratory measurements.

# Contents

<b>Supervisory Committee</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Table of Contents</b>	<b>v</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>x</b>
<b>Acknowledgements</b>	<b>xvi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Ground-based optical/near IR telescopes in the coming decades . . .	1
1.2 Principles of Adaptive Optics . . . . .	2
1.2.1 Classical Adaptive Optics . . . . .	2
1.2.2 Necessary evolution of AO technology . . . . .	3
1.2.3 Multi-Conjugate and Ground Layer AO . . . . .	3
1.2.4 Multi-Object AO . . . . .	4
1.2.5 Assessing the challenges of Open Loop AO . . . . .	6
1.3 Raven: An MOAO science and technology demonstrator . . . . .	7
1.4 Research Objectives . . . . .	8
1.5 Contents of dissertation . . . . .	9
1.6 List of acronyms and abbreviations . . . . .	10
<b>2 Adaptive Optics: Mathematical Fundamentals and Technologies</b>	<b>13</b>
2.1 Atmospheric turbulence . . . . .	13
2.1.1 Covariance and structure functions . . . . .	14
2.1.2 The index of refraction structure function . . . . .	16
2.1.3 Phase aberrations of wavefronts . . . . .	20

2.1.4	The effects of optical turbulence on imaging . . . . .	22
2.1.5	Angular anisoplanatism . . . . .	23
2.1.6	Laser Guide Stars . . . . .	24
2.1.7	Representing the phase . . . . .	25
2.2	Adaptive Optics technologies . . . . .	27
2.2.1	The Shack-Hartmann WFS . . . . .	28
2.2.2	Deformable mirrors . . . . .	29
2.2.3	Command estimation for SCAO . . . . .	30
2.2.4	Principles of WF estimation for WFAO . . . . .	31
2.3	Raven: A Multi-Object Adaptive Optics test bed . . . . .	32
2.4	Wavefront reconstruction modes on Raven . . . . .	34
2.4.1	Closed Loop Mode . . . . .	36
2.4.2	Ground Layer AO . . . . .	36
2.4.3	Static MOAO Mode . . . . .	36
2.4.4	MOAO with CL-WFS Mode . . . . .	36
2.4.5	MOAO with Figure WFS mode . . . . .	37
<b>3</b>	<b>Tomography for MOAO</b> . . . . .	<b>38</b>
3.1	Sky coverage . . . . .	38
3.2	Static Minimum Mean Square Error . . . . .	39
3.2.1	Explicit layered static MMSE . . . . .	40
3.2.2	Spatio-Angular static MMSE . . . . .	42
3.2.3	Equivalence of static MMSE reconstructors . . . . .	43
3.2.4	Overcoming temporal lag error . . . . .	44
3.2.5	Linear Quadratic Gaussian Controller . . . . .	46
3.3	Modal reconstruction . . . . .	50
3.3.1	Static Reconstructors in Zernike Space . . . . .	50
3.3.2	Predictive MMSE in Zernike space . . . . .	52
3.3.3	Theoretical temporal lag errors . . . . .	65
3.3.4	Linear Quadratic Gaussian controller with modal prediction models . . . . .	69
3.4	Zonal Reconstructors . . . . .	70
3.4.1	Zonal static MMSE reconstructors . . . . .	71
3.4.2	Predictive Spatio-Angular MMSE . . . . .	73
3.4.3	Bi-linear spline interpolation model . . . . .	77

3.4.4	Zonal Linear Quadratic Gaussian controllers . . . . .	77
<b>4</b>	<b>Results of Tomographic Wavefront Reconstruction Simulations of Raven</b>	<b>80</b>
4.1	End-to-end simulation . . . . .	81
4.1.1	Temporal properties . . . . .	81
4.1.2	Spatial properties . . . . .	82
4.1.3	WFS linearity . . . . .	82
4.1.4	WFS noise . . . . .	84
4.1.5	Science images . . . . .	84
4.2	Parameter space study . . . . .	86
4.3	Modal reconstructors results and analysis . . . . .	90
4.3.1	Static reconstructors: Verifying the equivalence of explicit and SA formulations . . . . .	90
4.3.2	Comparison of first and second order predictive models . . . . .	92
4.3.3	LQG results and analysis . . . . .	93
4.3.4	Meeting Raven science requirements . . . . .	95
4.4	Zonal basis results and analysis . . . . .	98
4.4.1	Static reconstructor . . . . .	99
4.4.2	Predictive reconstructor . . . . .	100
4.4.3	LQG . . . . .	102
4.5	Robustness . . . . .	103
4.5.1	Robustness to wind profile error . . . . .	103
4.5.2	Flat $C_n^2$ profile . . . . .	105
4.5.3	Over-modelled atmosphere . . . . .	106
4.6	MOAO with CL-WFS correction mode . . . . .	106
<b>5</b>	<b>Implementing tomographic reconstruction on Raven</b>	<b>108</b>
5.1	System calibration overview . . . . .	108
5.1.1	Retrieving reconstructor components from the system . . . . .	109
5.1.2	CU neutral density filter magnitudes . . . . .	110
5.1.3	Transformation between WFS measurement bases . . . . .	111
5.1.4	Identification of atmosphere parameters from the system . . . . .	116
5.2	Static reconstructor results . . . . .	116
5.2.1	Comparison of static MOAO methods . . . . .	118

5.2.2	Learning on the fly . . . . .	118
5.2.3	Comparing MOAO to other WF correction modes . . . . .	121
5.3	Dynamic reconstructor results . . . . .	123
5.3.1	RTC implementation of the LQG . . . . .	123
5.3.2	Results with WFS frame rates and guide star magnitudes . . .	125
5.4	Robustness to dynamic errors . . . . .	129
5.4.1	Sensitivity to atmospheric model layer altitude error . . . . .	130
5.4.2	Field rotation update constraints . . . . .	132
5.5	First on-sky results . . . . .	133
<b>6</b>	<b>Conclusions</b>	<b>137</b>
6.1	Future Work . . . . .	138
	<b>Bibliography</b>	<b>140</b>
<b>A</b>	<b>Performance Modeling for the RAVEN Multi-Object Adaptive Optics Demonstrator</b>	<b>148</b>
<b>B</b>	<b>Static and predictive tomographic reconstruction for wide-field multi-object adaptive optics systems</b>	<b>165</b>

# List of Tables

Table 2.1	Raven specifications. . . . .	35
Table 4.1	Subaru Model Profiles for Raven . . . . .	87
Table 4.2	Raven Baseline Configuration Parameters used in all modal simulation cases presented below. . . . .	91
Table 4.3	Raven End-to-End simulation results. The optimal performance ( % ensquared energy) for each GS magnitude is shown for three reconstructors: the static MMSE, Spatio-Angular prediction and the AR2 prediction model. AR1 prediction is included for comparison purposes . . . . .	94
Table 4.4	Raven End-to-End simulation results. The optimal performance ( % ensquared energy) for each GS magnitude is shown for the LQG reconstructor using the Spatio-Angular prediction model and the AR2 prediction model. . . . .	96
Table 4.5	Raven End-to-End simulation results. The optimal performance ( % ensquared energy) for each GS magnitude is shown for the zonal static SA, and compared to SA Prediction. . . . .	102
Table 4.6	Raven End-to-End simulation results. The optimal performance ( % ensquared energy) for each GS magnitude is shown for the zonal static SA, and compared to SA LQG. . . . .	103
Table 4.7	Sensitivity of reconstructors to error in the $C_n^2$ profile estimate. . . . .	106
Table 4.8	Projected improvement in strehl ratio and ensquared energy when slow CL-WFS measurements of the science target are periodically added as offsets to the real-time (500Hz) tomographic correction. . . . .	107
Table 5.1	Atmospheric parameters in three test cases . . . . .	122

# List of Figures

Figure 1.1 Atmospheric tomography: Portions of the atmosphere are measured in the GS directions, then estimated at discrete layers. The estimate in each layer is then cropped and propagated to the pupil in the science direction. . . . .	5
Figure 2.1 The Hufnagel $C_n^2$ model plotted by Valley [1] for various wind ratios. . . . .	18
Figure 2.2 Mt Graham $C_n^2$ profile measurements [2] . . . . .	18
Figure 2.3 Kolmogorov vs von Kármán PSDs with different outer scales [3].	19
Figure 2.4 Effects of atmospheric turbulence on a star (left), AO compensated image (right)[4]. . . . .	23
Figure 2.5 Visualization [5] of the first few radial orders of Zernike polynomials in the ordering given by Noll [6]. . . . .	26
Figure 2.6 Basic closed loop schematic for a single conjugate AO system. .	27
Figure 2.7 The effects of turbulence on SH-WFS spot positions. Plane waves (top) focus to diffraction limit at the centre of the lenslets; aberrated wavefronts (bottom) cause distortion in spot shape and overall displacement of the spot centres. . . . .	28
Figure 2.8 Functional optical block diagram of RAVEN. Dashed blocks are deployable. Raven consists of 8 main subsystems: the deployable Calibration Unit, the Open-Loop NGS WFSs, the Science Pick-offs, the Science Relays, the Closed-Loop NGS Truth/Figure WFSs, the Beam Combiner, the LGS WFS and the Acquisition Camera. . . . .	33
Figure 2.9 3D representation of Raven. Highlight on one CL-WFS path and figure source which will enable more WF correction modes. . .	37

Figure 3.1	Temporal diagrams. Top: $\text{Int}(\Delta/T_s) = 0$ , Commands $u_k$ are conditioned to measurement $\mathbf{s}_{k-1}$ . Bottom: $T_s < \Delta < 2T_s$ ; $\mathbf{u}_k$ is conditioned to $s_{k-2}$ . . . . .	45
Figure 3.2	Aliasing error in the computation of explicit layered covariance matrix, $\mathbf{P}_\alpha \Sigma_\varphi \mathbf{P}_\alpha^\top$ , is reduced with increasing radial order. . . . .	52
Figure 3.3	Parameter identification using different search minimization algorithms. BFGS search provides a speed up factor of more than 2 orders of magnitude in the AR2 case. . . . .	59
Figure 3.4	Bode plots of continuous and discrete TFs. . . . .	61
Figure 3.5	Theoretical temporal auto-correlation functions assuming frozen-flow against the 2nd-order continuous and discrete predictive models fitting the initial 50 <i>ms</i> of the theoretical curves. . . . .	63
Figure 3.6	Temporal auto and cross-correlation functions for the tip and focus modes. Although at $t = 0$ these plots show the spatial values in [6], for $t > 0$ correlations appear and vanish as shown. One can see, for the case of focus, that although it isn't correlated to any other mode plotted for $t = 0$ , a strong (anti) correlation appears with tilt as time elapses. . . . .	64
Figure 3.7	Comparison of temporal lag errors on an equivalent single layer atmosphere see Table 4.2 for further parameters. . . . .	67
Figure 3.8	Temporal lag error is traded by increased noise propagation through the wave-front reconstruction. Black markers: increase in limiting magnitude; blue markers: combined temporal plus noise propagation error; red circles indicate the increased limiting magnitudes for the minima of $\sigma$ . . . . .	69
Figure 3.9	Top: Coordinate system used to define the coordinates of phase point grid centre in an atmospheric layer at altitude $h$ . Bottom: Phase point locations: The ground grid (black diamonds) is expanded to fill the meta-pupil at layer $i$ (blue crosses) and the target phase point locations (red dots) must be interpolated onto the ground grid locations. . . . .	74
Figure 4.1	Atmosphere simulated phase temporal auto-corelation compared to theory: Does not match. . . . .	82

Figure 4.2 Atmosphere simulated phase temporal auto-correlation matches theory for any sample rate after code correction. . . . .	83
Figure 4.3 Atmosphere simulated phase spatial auto-correlation compared to theory: Good agreement. . . . .	83
Figure 4.4 Linearity test of the diffractive model of a SH WFS as a function of magnitude without threshold (left) and with threshold (right). . . . .	84
Figure 4.5 Simulated WFS noise properties and photon counts. . . . .	85
Figure 4.6 Baseline test configuration: A regular asterism with a radius of 45", performance was sampled at regular intervals from on-axis out to 60" in opposite directions. Average performance is computed from samples out to 30" regardless of asterism radius. . . . .	88
Figure 4.7 Performance for baseline system. Blue triangles are rms WFE from full error signal, green squares are T/T removed rms WFE (left axis). Open circles are EE and red X's are Strehl ratio (right axis). . . . .	88
Figure 4.8 Performance for asterisms of varying diameter: Green squares are WFE (left axis). Open circles are EE and red X's are Strehl ratio (right axis). . . . .	89
Figure 4.9 As the number of Zernike modes used to compute the explicit layer reconstructor increases, system performance approaches that of the spatio-angular reconstructor which is computed using an analytical expression and effectively accounts for an infinite number of modes. . . . .	92
Figure 4.10 Simulation results showing peak values of Strehl ratios (top) and ensquared energy (bottom) and the WFS framerate at-which they occur as a function of tomographic algorithm for each magnitude using stand alone prediction models AR1, AR2 and 1-step SA. . . . .	95
Figure 4.11 The simulated science images for the SA predictive model (left) and the static reconstructor (right) in modal space. . . . .	96
Figure 4.12 Static modal reconstructor is replotted against the modal LQG algorithm simulation results using the AR2 prediction model and the SA prediction model. . . . .	97

Figure 4.13	The use of more complex predictive algorithms will allow Raven to meet scientific requirements using dimmer NGSs therefore increasing sky coverage. . . . .	98
Figure 4.14	Static reconstructors in zonal and modal basis. A discrepancy in performance can be traced to modal aliasing. . . . .	99
Figure 4.15	Strehl ratios and ensquared energy as a function of algorithm and magnitude using the static zonal reconstructor vs the predictive zonal SA reconstructor. . . . .	101
Figure 4.16	Strehl ratios and ensquared energy as a function of algorithm and magnitude for all three static and dynamic zonal reconstructors. . . . .	104
Figure 4.18	The data pipeline when injecting command offsets to the MOAO correction from slow CL-WFS frames from the science targets . . . . .	107
Figure 5.1	Schematic for pinpointing the location of objects in the FoR. . . . .	109
Figure 5.2	Measured locations of pinholes in Raven calibration unit relating model space coordinates to system coordinates. . . . .	110
Figure 5.3	Source magnitudes provided by the CU as a function of pinhole location, filter and lamp power. . . . .	111
Figure 5.4	The OL-WFS axes are rotated with respect to each other and the CL-WFS axes. . . . .	112
Figure 5.5	Common disturbances which span the vector space are applied to the ground-conjugated CDM and measured by all WFSs. . . . .	114
Figure 5.6	A one-to-one reordering matrix to link slopes readout of the RTC to computations carried out in Matlab. . . . .	114
Figure 5.7	Left: simulated WFS frame of z11. Middle: slopes measured by RTC and Matlab. Right: slopes measured by RTC and Matlab with order transformation matrix applied to RTC slopes . . . . .	115

Figure 5.8 (a) Left: Expected CU atmosphere profile according to its design. Right: Measured CU atmosphere profile identified using on-board SLODAR. (b) Atmosphere identified during observation; results are compared to CFHT MASS measurements. There is no comparison for the ground layer as the CFHT data does not include a ground layer estimate. It would be different anyway due to Subaru dome seeing. The Raven profiler is constrained to estimate a maximum height of 12km and therefore attributes all turbulence from higher layers to that height. . . . .	117
Figure 5.10 Measured $\Sigma_{\alpha,\alpha}$ . . . . .	120
Figure 5.11 modeled $\Sigma_{\beta,\alpha}$ . . . . .	120
Figure 5.12 Comparison of a measured covariance matrix to a modelled covariance matrix. . . . .	121
Figure 5.13 Static explicit zonal reconstruction compared to GLAO and SCAO in a ground layer-dominated case. . . . .	122
Figure 5.14 Static explicit zonal reconstruction compared to GLAO and SCAO in a low total turbulence case. . . . .	123
Figure 5.15 Static explicit zonal reconstruction compared to GLAO and SCAO in a realistic atmosphere case. . . . .	124
Figure 5.16 J-Band Science Images of GLAO vs MOAO for a static zonal SA tomographic reconstructor. Asterism diameter was 2 arc min and NGS magnitudes were 11.25,12.13, and 12.13. . . . .	125
Figure 5.17 (a) LQG control scheme. (b) LQG control scheme as implemented in the RTC. . . . .	126
Figure 5.18 Asterism used in test cases. . . . .	127
Figure 5.19 Results using static, predictive and LQG zonal reconstructors at varying framerates for different GS magnitudes using the ground layer dominated atmosphere. . . . .	128
Figure 5.20 Performance comparison for Static, Predictive and LQG tomographic reconstruction algorithms with varying framerate using the distributed atmosphere profile. . . . .	129
Figure 5.21 PSFs of best performance science images for the static, predictive and LQG algorithms. . . . .	130

Figure 5.22	Variation of the single-layer atmosphere model used in the recon- structor shows the behaviour of the residual error as a function of model layer altitude. . . . .	131
Figure 5.23	Update of the layer altitude profile during an observation. . . . .	132
Figure 5.24	Update of the GS positions during an observation. . . . .	133
Figure 5.25	First on-sky results: Images from Raven science camera show performance for different AO modes [7]. . . . .	134
Figure 5.26	Observation field used for first on-sky MOAO test. . . . .	135
Figure 5.27	Saturn imaged on IRCS with and without MOAO correction [7].	136

## ACKNOWLEDGEMENTS

I would like to acknowledge the work of the Raven team as an essential element to the success of this work. I would also like to acknowledge the dedication and commitment of my mentor, Dr. Correia, who continues to advise me, even from another continent.

# Chapter 1

## Introduction

### 1.1 Ground-based optical/near IR telescopes in the coming decades

Ten meter class optical/near IR ground based telescopes are still producing ground-breaking science thanks to constant improvements and updates to their instrument packages. The next step in the evolution of telescope technology will be the Extremely Large Telescope (ELT) class of telescopes. These have the potential to revolutionize our view of the universe. The Thirty Meter Telescope (TMT) [8] the European ELT (EELT) [9] and the Giant Magellan Telescope (GMT) [10] are all approaching their construction phases. Some first light instruments on these ELTs will use seeing limited light, while others will be fed by facility Adaptive Optics (AO) systems. These facility systems are a first pass at Wide Field AO (WFAO) technology on such a large scale; choices regarding the amount of risk to take on have been made which limit the kind of instruments they will be able to feed, and the amount of multiplexing that can be carried out for certain science cases.

The second generation of ELT instruments are also being planned; two of the major ELT science cases, characterizing first light objects in the universe and performing detailed measurements of galaxies in the era of peak star formation, are strongly driving the development of a specific kind of WFAO instrument, called Multi-Object AO (MOAO). For example, near IR spectrographs with greater than 20 deployable Integral Field Units (IFUs) over a 5 to 10 arc-minute field of regard (FoR) are highly desirable potential instruments on ELTs because they are ideally suited for studying the evolution of galaxies from first light to the era of peak star formation. This will

only be possible, however, when these instruments are fed by an MOAO system.

## 1.2 Principles of Adaptive Optics

AO is a well established technique of compensating for atmospheric turbulence that blurs the light from astronomical objects and degrades the resolution capabilities of a telescope with diameter  $D$  from its diffraction limit ( $1.22\lambda/D$ ) to the seeing limit ( $\lambda/r_0$ ). The effect is to scale the resolution limit of a large telescope to that of a much smaller telescope. The parameter  $r_0$  is called the Fried parameter; it is a quantification of the turbulence strength (Sec. 2.1.3). The effects of the atmosphere can be mitigated by placing telescopes at strategic locations high on mountain tops, such as Mauna Kea and Haleakela in Hawaii and on high desert plains such as Armazones and Atacama in the Chilean Andes. This serves both to reduce the amount of atmosphere above the telescope and reduce the humidity in the ground layer, but there is still a large amount of image quality that can be recovered by AO.

### 1.2.1 Classical Adaptive Optics

The most basic AO system design is a single conjugate (SC) closed loop system. Control is executed by a simple integrator with a reasonable gain. There are three main components in any AO system: the wavefront sensor (WFS), deformable mirror (DM) and the control system. Observing a reference source with a WFS, the estimated integrated wavefront distortion introduced to a flat wavefront (WF) travelling from the top of the atmosphere to the telescope can be used to command the DM to optically compensate for this error.

A severe limitation of classical AO is anisoplanatism which is the result of viewing two sources through the atmosphere separated by a given angular distance (Sec. 2.1.5). The two sources are typically a Guide Star (GS) and the science object of interest. The GS is a reference source used for WFSing; it is necessary, among other reasons, because the science target is most often not bright enough or compact enough to act as a good reference, a near-by bright star must be used. Because AO works by estimating the turbulence in the column of atmosphere above the telescope in the direction of a GS and then using that estimate to correct the image in the direction of a science object, there is always error corresponding to the different paths the wavefronts travel through the atmosphere. The characteristics of the atmosphere at

any given moment determine the size of the isoplanatic patch - the area on the sky within which the atmosphere behaves in roughly the same manner, typically about 20 arc-seconds. Outside of this patch, correction quality drops off sharply.

### 1.2.2 Necessary evolution of AO technology

The current generation of large optical telescopes and future ELTs are becoming increasingly expensive to operate on an hourly basis. As such, multiplexing observations to simultaneously take advantage of many interesting objects within the FoR is a necessary step. AO plays a key part in a variety of science goals and its utilization in future wide field instruments is required. Ways to extend the high quality of correction delivered by classical closed loop SCAO within the isoplanatic patch over more area of the sky, without requiring a GS within the isoplanatic area of each target, must be implemented. This is the goal of WFAO. There are several solutions to this problem, but the most promising methods require a more sophisticated knowledge of the real time properties of the atmosphere during observation.

Each method presents challenges, of which some are general to WFAO and some specific to an individual method. Each challenge must be addressed by way of simulation, laboratory tests and small-scale pathfinder instruments in order to reduce the risks of developing such an instrument on the scale of an ELT facility-class instrument.

### 1.2.3 Multi-Conjugate and Ground Layer AO

A first order solution to the WFAO problem is Ground Layer AO (GLAO) which makes the assumption that the majority of the WF phase aberrations are produced by turbulence located close to the ground. There are two additional approaches to WFAO which require a tomographic estimation of the three dimensional atmospheric wave-front disturbance above the telescope. Tomography, in the context of AO, is a back-projection of a very limited number of views using a large amount of *a priori* assumptions about atmospheric parameters derived either from models or other sets of measurements, for example, SLOpe Detection And Ranging (SLODAR) [11]. Making an estimate of the turbulent volume requires information from multiple WFSs locked on multiple GSs that probe different lines of sight through the atmosphere. A correction must then be computed from this estimate and there are different ways in which this can be done. The first approach is to place multiple DMs in series, each conjugated to a different atmospheric altitude. Atmospheric tomography was

conceived as a method to measure the instantaneous phase perturbations in the atmosphere to deduce the control signals for these DMs in what is called Multi-Conjugate AO (MCAO) [12, 13]. This MCAO [14, 15, 16] approach can be used to enlarge the Field of View (FoV) to sizes of an arc-minute or two, but the performance will ultimately still be limited by generalized anisoplanatism [17]. The FoV can be further enlarged by adding even more DMs in series, to remove the turbulence generated at even more atmospheric heights, but the complexity of the MCAO system rises (and the throughput falls) with each additional DM relay.

### 1.2.4 Multi-Object AO

MOAO is a parallel approach that promises to increase the field over which AO corrections can be applied to 5 or even 10 arc-minutes [18]. MOAO systems use the fact that there are only so many interesting targets in a given FoR. If a sufficiently accurate measurement is made of the turbulent volume over a telescope, one can place a probe with an embedded DM anywhere in the FoR and make the optimal turbulence correction for that position. To achieve the multiplexing advantage over a large FoR, MOAO systems use several DMs. Light from individual scientific targets are directed into separate optical paths, each containing a DM thus enabling simultaneous multi-wave-front correction. In essence, this means that for a large FoR, the turbulence-induced aberrations are compensated only within a few smaller fields. This simultaneous correction from a set of measurements implies that MOAO systems must run open loop, i.e., the WFSs are separated physically from the DM optical paths.

Making a measurement on each of several GSs within the FoR provides the requisite lines-of-sight through the atmosphere. Once the information from these multiple WFSs is combined into a single tomographic model of the turbulence [13], it is straight-forward to imagine multiple science pick-offs in parallel, each incorporating its own DM, feeding multiple IFUs or being directed to an imaging instrument. Falcon for VLT was the first proposed MOAO Integral Field Spectrograph (IFS) [18, 19], and it has served as a model for the more recent IRMOS and Eagle studies which are instrument concepts proposed for ELTs. Falcon was proposed as a WFAO science instrument to do 3D spectroscopy of IR galaxies - a process which requires high resolution over a wide FoV. They proposed using “several independent AO systems spread in the focal plane” with 3 WFSs per IFU measuring the off-axis wavefront

coming from stars located around a galaxy. The on-axis wavefront from the galaxy would be deduced from the off-axis measurements via tomography and corrected with the AO system within each IFU. The process of on-axis wavefront reconstruction from off-axis measurements was to be repeated as many times as there are spectroscopic IFUs. This instrument was proposed to have a 25 arc-min FoR with each IFU “local AO system” FoV having a 3 arc-min FoR [19].

In the modern MOAO case, a tomographic estimate of the phase in the volume is made from Open Loop (OL)-WFS measurements and then the total integrated phase in the pupil in an independent direction within the FoR is estimated (Fig. 1.1).

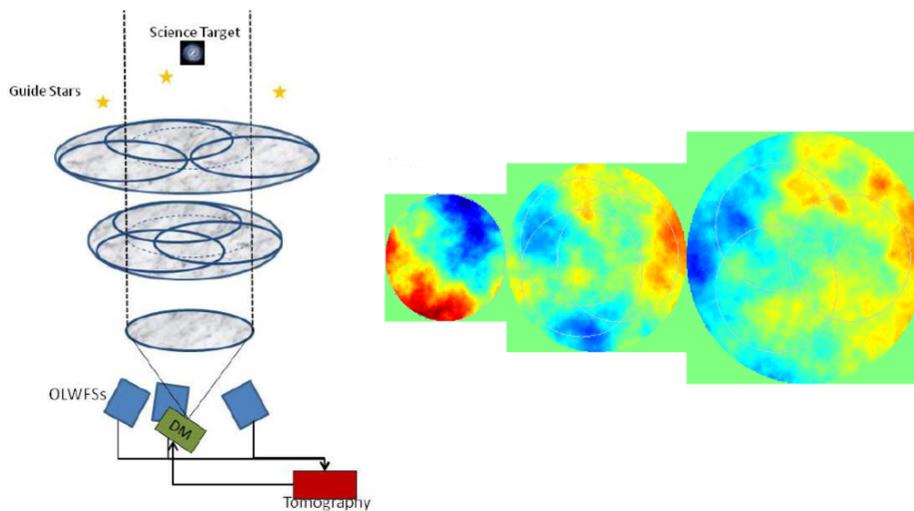


Figure 1.1: Atmospheric tomography: Portions of the atmosphere are measured in the GS directions, then estimated at discrete layers. The estimate in each layer is then cropped and propagated to the pupil in the science direction.

The multiplexing advantage of 20-arm ELT MOAO IFUs will enable fundamentally new science. The large surveys of 100’s of galaxies and first light objects needed to better understand these eras are not possible without multiplexing because the telescope time required to do these surveys with a single IFU would be highly prohibitive.

### 1.2.5 Assessing the challenges of Open Loop AO

Open loop AO systems are perceived as one of the major risks of building MOAO-fed instruments. Some other major questions include: How does one calibrate a MOAO system? How does one keep a MOAO system in alignment during the course of regular operations? Are there limits in the performance of an open loop AO system? How can we reliably trust that the DM is taking the proper shape and is being well-controlled?

Many of the challenges involved in designing a MOAO system, such as the use of tomography, Micro-Electro Mechanical Systems (MEMS) mirrors, and woofer-tweeter control, have all been demonstrated to work in different lab settings and are included in advanced instrument concepts. OL control, however, is perhaps the greatest risk to MOAO, partly because it is the biggest unknown. In an AO system with OL control, the WFSs do not sense the correction applied by the DM. Instead, the WFSs sense the full turbulent phase of the atmosphere and the DMs are commanded to take the appropriate shape without benefit of any feedback. While OL control is not a new idea, so-called “go to” adaptive optics was first used to make corrections and take science images immediately following pulses from a Laser Guide Star (LGS) (see Sec. 2.1.6) with a low duty cycle[20], interest in implementing open loop control on sky has been re-invigorated in the past few years. OL control introduces unique requirements on an AO system: the WFS needs to have a high dynamic range; effects of DM hysteresis and non-linearity need to be mitigated; finally, alignment and calibration become more challenging.

A considerable amount of work both theoretical and practical with respect to WFAO and tomography has been undertaken in the past two decades. More recently, dedicated efforts toward proofing the concepts and mitigating the risks of MOAO have been made on multiple fronts, starting from OL-AO demonstrators and working up to a single science channel MOAO testbed. Work has been done in both the MOAO and MCAO context on improving tomographic reconstruction techniques, specifically in low Signal to Noise Ratio (SNR) regimes to improve performance on dim Natural GSs (NGSs), thus increasing sky coverage. Extensive work is also being carried out on computational efficiency in anticipation of ELT WFAO instruments which will see orders of magnitude increases in the sizes of the matrices which will need to be manipulated within the data processing pipeline [21].

While the risks associated with MOAO IFSs have kept proposed Very Large Telescope (VLT) and ELT instruments on the drawing board, the scientific promise is

so great that multiple on-sky demonstrators have been developed. The Visible Light Laser Guide star Experiments (ViLLaGEs) [22] is a MEMS DM-based AO testbed on the Nickel 1-meter telescope at Lick Observatory. ViLLaGEs carried out on-axis experiments in both closed and open loop with NGSs and LGSs. It was the first on-sky experiment to successfully demonstrate open loop control. ViLLaGEs is one of the test beds that is being employed to develop the Keck Next Generation Adaptive Optics (NGAO) [23, 24] instrument, which is a tomographic, high-order open-loop AO system.

The Victoria Open Loop Test-bed (VOLT) [25] was an experiment aimed at distilling the problems of open loop control into a simple experiment. VOLT demonstrated open loop control in the laboratory and on-sky at the Dominion Astrophysical Observatory 1.2 m telescope using a simple, on-axis NGS system [26]. Both the VOLT and ViLLaGEs open loop AO demonstrators performed below expectations at low temporal frequencies, which seems to indicate small misalignments in open loop AO systems may ultimately limit their performance. Neither ViLLaGEs nor VOLT carried out on-axis wavefront estimation from off-axis sources (tomography) but are pivotal to MOAO instrument development as these experiences have led to a second generation of MOAO demonstrators that emphasize both calibration and alignment techniques.

Canary is a MOAO demonstrator at the William Herschel Telescope [27, 28, 29] that is considered a pathfinder for Eagle (recently renamed MOSAIC) on the EELT. The goals of the Canary project are to perform NGS, and subsequently LGS, based tomographic wavefront sensing, perform open-loop AO correction on-sky, and develop calibration and alignment techniques. This experiment saw first light in the Fall 2010 and achieved a MOAO Strehl of 26% (in Hband) [30]. They have since successfully carried out single channel MOAO science, [31] and on-sky testing of laser tomographic AO (LTAO) and uses 3 NGSs and 4 LGSs [32].

### **1.3 Raven: An MOAO science and technology demonstrator**

Raven will be the first MOAO instrument on an 8 m class telescope feeding an AO-optimized science instrument: the Subaru InfraRed Camera and Spectrograph (IRCS). The instrument has been designed and constructed at the University of Victoria Adaptive Optics Laboratory and shipped to Mauna Kea. It is stored on the

Cassagrain floor and is periodically installed on the IR Nasmyth platform of the 8.2m Subaru telescope, in between the telescope Nasmyth focus and IRCS [33]. Raven features 3 NGS WFSs, 1 LGS WFS and 2 science channels patrolling a 3.5' diameter FoR. The 2 science channels are each fed by a deployable science pick-off mirror, as envisioned for MOSAIC and IRMOS.

## 1.4 Research Objectives

The main challenges of an MOAO system have been identified as: Developing calibration procedures, computing the tomographic reconstructor, and designing the pick-off systems. Of these challenges, tomographic reconstructors for an MOAO system is the initial focus of this thesis. This work was carried out in the context of the Raven project and has attempted to achieve the following specific research objectives:

- Develop an innovative approach to MOAO tomography and be the first to implement it on a science-capable instrument.
- Design a unique, Raven-specific, wavefront reconstruction pipeline which utilizes the characteristics of the opto-mechanical design to augment performance by using all available measurement data.
- Conduct a comparative study of the unique Raven-specific algorithms vs established AO methods to establish the most effective approach both for the instrument, and in the general case of MOAO.
- Show definitively in the lab and on-sky that MOAO provides superior correction than GLAO.
- Implement a fully functional tomographic reconstructor on the real time computer (RTC) of Raven such that the minimum performance requirements of the project are met.
- Improve the performance of Raven in the laboratory environment using the telescope simulator. Improved performance refers both to increased sky coverage

and improved image quality; building on the experience of other projects to better understand the sources of errors in a MOAO system is key to developing novel approaches to attempt to mitigate them.

The contribution of this work to the AO research community is on multiple fronts. Foremost, is the development of innovative methods of implementation of tomographic wavefront reconstruction for MOAO including: A so-called spatio-angular (SA) linear quadratic gaussian (LQG) algorithm which scales in complexity as the order and number of system WFSs rather than the number and altitude of atmospheric layers. Second is the handling within the LQG algorithm of the asynchronous case, in which total system delay is not equal to an integer number of frames. Finally, it also makes a pivotal contribution as an individual part of the overall instrument project which is the first attempt to construct a MOAO system with more than a single science channel; it will be the first MOAO instrument to be installed on an 8m class telescope which will enable scientific observations to be made on such a system for the first time. Eventually this technology will lead to astronomical observations which have never before been possible.

The significance of this work in developing novel wavefront reconstruction algorithms includes the vital contribution it will make to the AO community's knowledge and understanding of the challenges and risks faced by MOAO for large telescopes. By applying the work to a pathfinder instrument project, it progresses the AO community toward MOAO for ELTs as well as provides a first opportunity for astronomers to propose and carry out science observations on an MOAO instrument on an 8m class telescope. The process of wavefront reconstruction is central to the success of the Raven project and therefore directly contributes to the future and advancement of MOAO instrumentation.

## 1.5 Contents of dissertation

**Chapter 1** of this dissertation has outlined some of the different AO methods and specified the distinct reasons for developing WFAO instruments along with the particular difficulties and challenges associated with doing so.

**Chapter 2** enumerates some basic concepts and technology required to do AO in general. It includes a brief description of several of the mathematical and statisti-

cal tools widely used to describe atmospheric turbulence and control WF correction instrumentation. Finally, there is a detailed description of the Raven project.

**Chapter 3** is a detailed development of the theoretical mathematical tools used to reconstruct the WF on an MOAO system. The algorithms progress from a generalized description of standard static tomographic reconstructor and proceed to dynamic reconstructors of varying complexity. These general algorithms are then developed specifically in a modal basis (Zernikes) and a zonal basis (phase points) with practical implementation in mind.

**Chapter 4** contains the results of Monte Carlo simulations of the end-to-end Raven system using each of the algorithms developed in Chapter 3. The performance of the methods is compared and the potential of the more complex algorithms to improve image quality and/or increase limiting magnitude over the baseline case is assessed. The complexity and hence practicality of implementation on the RTC and Raven Parameter Generator (RPG) is evaluated.

**Chapter 5** based on the evaluation and performance comparison carried out in simulation in Chapter 4, a cross section of reconstruction algorithms were implemented on the Raven system and tested in the laboratory environment using the telescope simulator. This chapter summarizes the results of these tests and makes a case for on-sky tests based on their potential to improve performance.

**Chapter 6** Conclusions and Future Work.

## 1.6 List of acronyms and abbreviations

<b>IR</b>	Infra Red
<b>ELT</b>	Extremely Large Telescope
<b>TMT</b>	Thirty Meter Telescope
<b>EELT</b>	European Extremely Large Telescope
<b>GMT</b>	Giant Magellan Telescope
<b>AO</b>	Adaptive Optics
<b>WFAO</b>	Wide Field Adaptive Optics
<b>MOAO</b>	Multi-Object Adaptive Optics
<b>IFU</b>	Integral Field Unit
<b>FoR</b>	Field of Regard
<b>WFS</b>	Wavefront Sensor

<b>DM</b>	Deformable Mirror
<b>WF</b>	Wavefront
<b>GS</b>	Guide Star
<b>SCAO</b>	Single Conjugate Adaptive Optics
<b>GLAO</b>	Ground Layer Adaptive Optics
<b>SLODAR</b>	SLOpe Detection And Ranging
<b>MCAO</b>	Multi-Conjugate Adaptive Optics
<b>FoV</b>	Field of View
<b>IFS</b>	Integral Field Spectrograph
<b>OL-WFS</b>	Open Loop Wavefront Sensor
<b>MEMS</b>	Micro-Electro Mechanical System
<b>SNR</b>	Signal to Noise Ratio
<b>NGS</b>	Natural Guide Star
<b>ViLLaGEs</b>	Visible Light Laser Guide Star Experiments
<b>LGS</b>	Laser Guide Star
<b>NGAO</b>	Next Generation Adaptive Optics
<b>VOLT</b>	Victoria Open Loop Test bed
<b>LTAO</b>	Laser Tomographic Adaptive Optics
<b>IRCS</b>	Infra Red Camera Spectrograph
<b>RTC</b>	Real Time Computer
<b>SA</b>	Spatio-Angular
<b>LQG</b>	Linear Quadratic Gaussian
<b>RPG</b>	Raven Parameter Generator
<b>PSD</b>	Power Spectral Density
<b>CFHT</b>	Canada France Hawaii Telescope
<b>SR</b>	Strehl Ratio
<b>FWHM</b>	Full Width at Half Maximum
<b>PSF</b>	Point Spread Function
<b>EE</b>	Ensquared Energy
<b>CCD</b>	Charged Coupled Device
<b>GeMS</b>	Gemini MCAO System
<b>T/T</b>	Tip/Tilt
<b>SH-WFS</b>	Shack-Hartmann Wavefront Sensor
<b>CoG</b>	Centre of Gravity
<b>SVD</b>	Singular Value Decomposition

<b>CU</b>	Calibration Unit
<b>CL-WFS</b>	Closed Loop Wavefront Sensor
<b>MMSE</b>	Minimum Mean Square Error
<b>AR</b>	Auto Regressive
<b>BFGS</b>	Broyden-Fletcher-Goldfarb and Shanno
<b>SDM</b>	Science Deformable Mirror
<b>CDM</b>	Calibration Deformable Mirror
<b>OOMAO</b>	Object Oriented Matlab Adaptive Optics
<b>MAOS</b>	Multithreaded Adaptive Optics Simulator
<b>WFE</b>	Wavefront Error
<b>IQ</b>	Image Quality
<b>NCPA</b>	Non-Common Path Aberration

## Chapter 2

# Adaptive Optics: Mathematical Fundamentals and Technologies

The successful operation of an AO instrument hinges upon several key elements: a well characterized system, including WF sensors, WF correctors, and a good calibration plan that can generate a reliable and robust relationship between them. A control loop with real-time and background components is required whether the system is open loop or closed loop. Particularly important in the case of WFAO and open loop control, is prior knowledge about the behaviour of the atmosphere. This chapter provides background information regarding atmospheric statistics. All of the tomographic WF reconstructors developed in subsequent chapters are based on these statistics. A brief introduction to the various types of WF sensors and correctors is provided and finally, the context of the work itself is established with a description of Raven, the MOAO science and technology demonstrator, and the way it will be used to achieve the research objectives outlined in Chapter 1.

### 2.1 Atmospheric turbulence

In the atmosphere, it is understood that small variations in temperature occur, which induce local changes in the wind velocity; the results of these two phenomena are evolving random irregularities in the atmosphere's index of refraction. It is this optical turbulence that is referred to as "atmospheric turbulence" throughout this dissertation (not to be confused with mechanical turbulence due to wind). This optical turbulence leads to wavelength-dependant degradation of image quality in an

optical/near IR imaging instrument, degrading the resolution of the detector from the diffraction limit,  $1.22\lambda/D$  where  $D$  is the aperture diameter of the telescope, to the so-called seeing limit,  $\lambda/r_0$  where  $r_0 \ll D$  ( $r_0$  is called the Fried parameter and is described in sec. 2.1.3). The main atmospheric effects on electromagnetic wavefronts are focusing and spreading, as well as scintillation (intensity changes), beam wander and speckles. Kolmogorov described turbulence in a fluid as the cascade of kinetic energy from large to small length scales [34]. He states that kinetic energy associated with large eddies in a fluid is redistributed without loss to successively smaller eddies until it reaches a regime below a given length scale called the inner scale  $I_0$ . Below this length scale, the energy is dissipated by small scale processes such as heat transfer. Because the fluctuations that cause the eddies are random, the atmosphere can be considered a locally isotropic, continuous random field and can therefore be characterized by some useful functions. The outer scale,  $L_0$ , is a (non-constant) radius outside of which the atmosphere can no longer be considered isotropic. It can range from a few meters to many tens of meters.

### 2.1.1 Covariance and structure functions

The covariance function,  $\mathcal{C}_f(\mathbf{r}_1, \mathbf{r}_2)$ , of a quantity which varies randomly with time or space,  $f(\mathbf{r})$ , characterizes the mutual relationship between the fluctuation of  $f(\mathbf{r})$  at location  $\mathbf{r}_1$  and  $\mathbf{r}_2$  [35]. An important property is homogeneity; a random field is called homogeneous if its ensemble average,  $\langle f(\mathbf{r}) \rangle$  is constant in  $\mathbf{r}$ . If this is the case, the covariance function depends only on the difference  $r = |\mathbf{r}_1 - \mathbf{r}_2|$ , that is,

$$\mathcal{C}_f(\mathbf{r}_1, \mathbf{r}_2) = \mathcal{C}_f(|\mathbf{r}_1 - \mathbf{r}_2|) = \mathcal{C}_f(|\mathbf{r}_2 - \mathbf{r}_1|). \quad (2.1)$$

The mean values of meteorological variables undergo slow, smooth changes and cannot therefore be considered strictly homogeneous in either the temporal or spatial domain. They can be described more generally using their structure function [34],[35]: if  $\langle f(r) \rangle$  is non-homogeneous, a new function is formulated from the difference,

$$F_\delta(r) = f(r + \delta) - f(r). \quad (2.2)$$

For values of  $\delta$  which are not too large (smaller than  $L_0$  for example), low frequency changes in  $f(r)$  can be neglected and  $F_\delta(r)$  is approximately homogeneous. Both spatial and temporal structure functions can be developed from the following steps:

Consider the difference function,  $F_\delta(r)$ , the covariance of that function is given by,

$$\mathcal{C}_F(r_1, r_2) = \langle F_\delta(r_1)F_\delta(r_2) \rangle. \quad (2.3)$$

By substituting Eq. 2.2 into Eq. 2.3 and expanding, it becomes clear that  $\mathcal{C}_F$  is composed of a linear combination of elements of the form,

$$D_f(r_i, r_j) = \langle [f(r_i) - f(r_j)]^2 \rangle, \quad (2.4)$$

which is called the structure function. In order for  $\mathcal{C}_F$  to be homogeneous (depend only on  $|r_1 - r_2|$ ), it is sufficient to use

$$D_f(|r_1 - r_2|) = D_f(\delta) = \langle [f(r + \delta) - f(r)]^2 \rangle. \quad (2.5)$$

This holds for any quantity,  $f$  changing randomly with time,  $t$ , or with spatial position,  $\mathbf{r}$ . If  $f(r)$  also has zero mean, expanding  $D_f(r)$  gives,

$$D_f(\delta) = \langle f(r + \delta)^2 \rangle + \langle f(r)^2 \rangle - 2\langle f(r + \delta)f(r) \rangle. \quad (2.6)$$

From stationarity, we can say that

$$\begin{aligned} \langle f(r)^2 \rangle &= \langle f(r)f(r) \rangle \\ &= \langle f(r + \delta)f(r + \delta) \rangle \\ &= \mathcal{C}_F(0), \end{aligned} \quad (2.7)$$

and

$$\langle f(r + \delta)f(r) \rangle = \mathcal{C}_F(\delta). \quad (2.8)$$

Using these definitions, the structure function can be expressed in terms of the covariance function,

$$D_f(\delta) = 2(\mathcal{C}_F(0) - \mathcal{C}_F(\delta)). \quad (2.9)$$

Next, make the realistic assumption that  $\mathcal{C}_F(\infty) \rightarrow 0$ , meaning that as the quantity of interest goes to infinity, the correlation between two occurrences of the random variable becomes negligible,

$$D_f(\infty) = 2\mathcal{C}_F(0). \quad (2.10)$$

Therefore, plugging this into Eq. 2.9 and rearranging, the covariance function can be

expressed in terms of the structure function,

$$\mathcal{C}_F(\delta) = \frac{1}{2}[D_f(\infty) - D_f(\delta)]. \quad (2.11)$$

The covariance of a homogeneous random field can also be expressed in terms of its power spectral density (PSD) function,  $\mathbf{W}(\boldsymbol{\omega})$ , by taking its inverse Fourier transform [36],

$$\mathcal{C}_F(\delta) = \int_{-\infty}^{\infty} e^{i\boldsymbol{\omega}\delta} \mathbf{W}(\boldsymbol{\omega}) d\boldsymbol{\omega}, \quad (2.12)$$

which can be rewritten,

$$\mathcal{C}_F(\delta) = \int_{-\infty}^{\infty} \cos(\boldsymbol{\omega}\delta) \mathbf{W}(\boldsymbol{\omega}) d\boldsymbol{\omega}, \quad (2.13)$$

because the covariance function is a real, symmetric function, i.e.  $\mathcal{C}_F(\delta) = \mathcal{C}_F(-\delta)$ . It can now be seen that the structure function of a homogeneous random process can be determined directly from its PSD [37],

$$D_f(\delta) = 2 \int_{-\infty}^{\infty} [1 - \cos(\boldsymbol{\omega}\delta)] \mathbf{W}(\boldsymbol{\omega}) d\boldsymbol{\omega}. \quad (2.14)$$

This holds true so long as the integral,

$$\int_{-\infty}^{\infty} \mathbf{W}(\boldsymbol{\omega}) d\boldsymbol{\omega} \quad (2.15)$$

exists, meaning the total power of the property must be finite.

### 2.1.2 The index of refraction structure function

It has been shown [38],[39],[35] that a relationship exists between velocity fluctuations in a turbulent flow and concentration fluctuations of a conservative passive additive, a property of the fluid element that does not change when the volume element is shifted in space (conservative), and whose quantity does not affect the dynamical regime of the turbulence, that is it does not exchange energy with the turbulence (passive). This leads directly to the so-called two-thirds law which states that the structure function of such a property,  $n$ , is expressed as,

$$D_n(r) = \langle [n(r_0) - n(r_0 + r)]^2 \rangle. \quad (2.16)$$

The expression for  $D_n$  can be derived [35] from the velocity structure function given by Kolmogorov [34],

$$D_v(r) = \langle [\mathbf{v}_r(r_0 + r) - \mathbf{v}_r(r)]^2 \rangle \quad (2.17)$$

assuming the atmosphere is locally homogeneous, locally isotropic and the turbulence is incompressible. For separations,  $r$ , sufficiently small, a 2/3 power law with  $r$  is observed:

$$D_v = C_v^2 r^{2/3} \quad (2.18)$$

$$D_n = C_n^2 r^{2/3} \quad (2.19)$$

The parameters  $C_v^2, C_n^2$  are structure constants.  $C_n^2$  is called the index of refraction structure constant and is a measure of the strength and distribution of the optical turbulence. This strength varies with altitude and with time; a commonly used model of the  $C_n^2$  profile is the Hufnagel-Valley model, an empirical model based on observations initially formulated by Hufnagel [40], and modified by Valley [1]:

$$C_n^2(h) = A \left[ 2.2 * 10^{-23} h^{10} e^{-h} \left( \frac{V_w}{\bar{V}_w} \right)^2 + 10^{-16} e^{-h/1.5} \right], \quad (2.20)$$

which has units of  $m^{-2/3}$ . The value  $A$  is a scaling constant,  $h$  is the vertical height and  $V_w/\bar{V}_w$  is the ratio of upper atmospheric wind speed to the mean upper atmospheric wind speed. The model for various wind velocity ratios are plotted by Valley and shown in Fig. 2.1. These can be compared to measurements of the  $C_n^2$  profile, for example those taken at Mt. Graham [2] are shown in Fig. 2.2. With a WFAO instrument, direct estimation of the current  $C_n^2$  profile from a near-by instrument or from on-board measurements will be required. The power law in Eq. 2.19 holds so long as the distance,  $r$ , remains smaller than the turbulence outer scale,  $L_0$ . According to the relationship between the structure function and PSD (Eq. 2.14), an expression for the spatial PSD can be derived from,

$$D_n(r) = 2 \int_{-\infty}^{\infty} [1 - \cos(\omega r)] \mathbf{W}(\omega) d\omega = C_n^2 r^{2/3}. \quad (2.21)$$

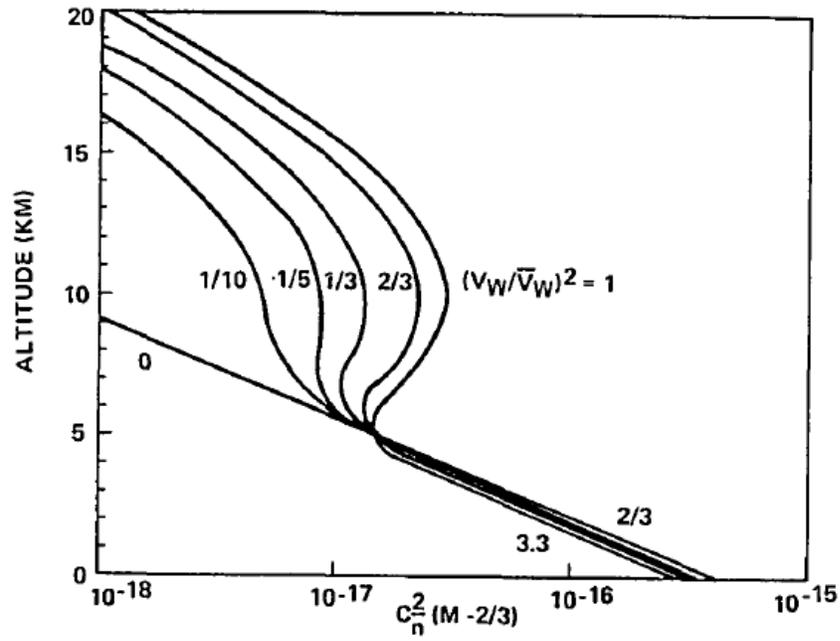


Figure 2.1: The Hufnagel  $C_n^2$  model plotted by Valley [1] for various wind ratios.

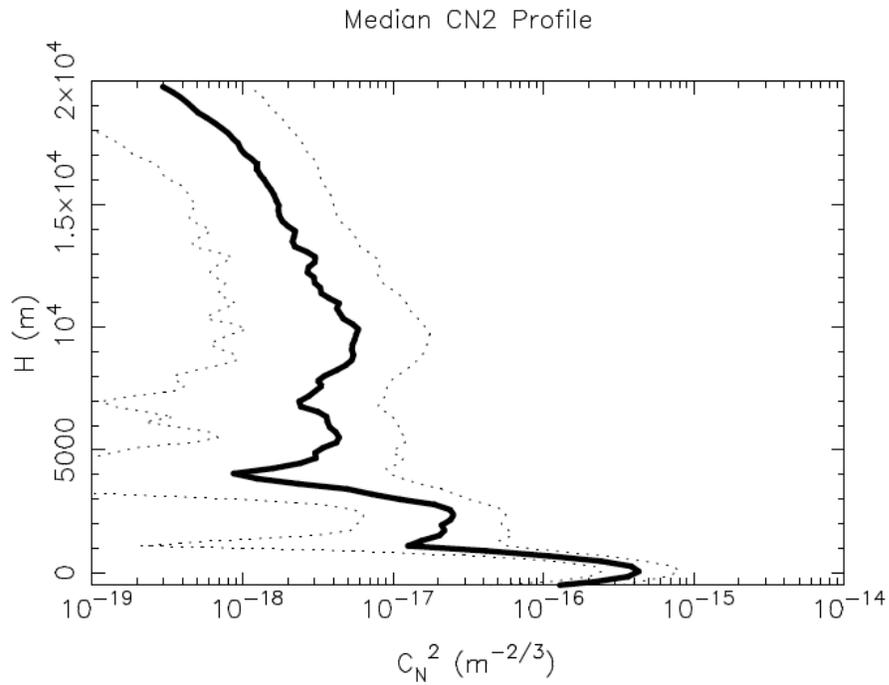


Figure 2.2: Mt Graham  $C_n^2$  profile measurements [2]

Solving for  $\mathbf{W}(\omega)$  ultimately results in the following expression for the index of refraction PSD [41],

$$\mathbf{W}(\omega) = \frac{5}{18\pi} C_n^2 \omega^{-3} \int_{I_0}^{L_0} \sin(\omega r) r^{-1/3} dr, \quad (2.22)$$

recalling that  $I_0$  is the inner-scale, the distance scale on-which small scale energy dissipation processes begin to take over. When the limits of integration are taken from 0 to infinity, this results in the Kolmogorov spectrum,

$$\mathbf{W}(\omega) = 0.033 C_n^2 \omega^{-11/3}. \quad (2.23)$$

For finite outer scales, the von Kármán spectrum eliminates the problem of infinite energy in the PSD as frequency goes to zero,

$$\mathbf{W}(\omega) = \frac{\Gamma(11/6)\pi^{-9/2}}{8\Gamma(1/3)} \left( \frac{C_n^2}{1.9\omega_0^{2/3}} \right) L_0^3 \left( 1 + \frac{\omega^2}{\omega_0^2} \right)^{-11/6}, \quad (2.24)$$

where  $\omega_0 = 2\pi/L_0$ . Fig. 2.3 shows both the Kolmogorov and von Kármán spectra for a finite outer scale.

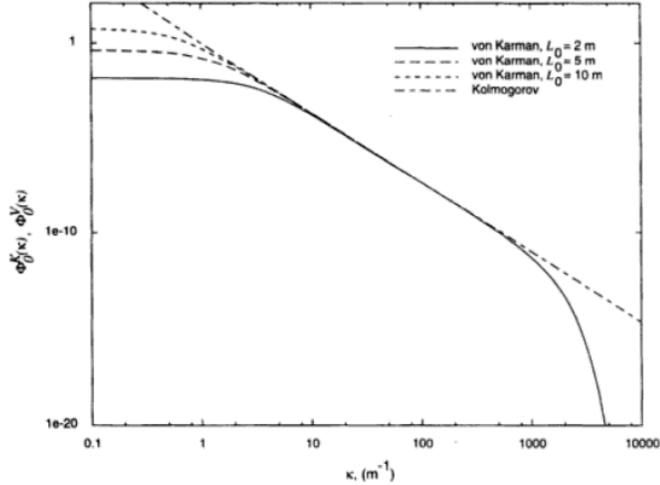


Figure 2.3: Kolmogorov vs von Kármán PSDs with different outer scales [3].

### 2.1.3 Phase aberrations of wavefronts

An optical wavefront is defined as a surface over-which the phase of a light wave has the same value. The light wave is generally expressed as a complex function with amplitude,  $A$  and phase,  $\varphi$ ,

$$\Psi = Ae^{i\varphi}. \quad (2.25)$$

Aberrations in the wavefront phase are introduced when a light wave passes through the distribution of index of refraction inhomogeneities in the atmosphere. The fluctuations are related to the index of refraction via,

$$\phi = k \int n(h)dh, \quad (2.26)$$

where  $k$  is the wavenumber of the light waves ( $2\pi/\lambda$ ) and the integral over  $h$  represents the total phase aberrations incurred by travel from the top of the atmosphere to the ground along the direction of travel. Throughout this work, the integrated phase at the ground is denoted by  $\phi$  and the distributed phase at an arbitrary height within the atmosphere by  $\varphi$ . The phase structure function,

$$D_\phi(\mathbf{r}) = \langle [\phi(\mathbf{r}_1) - \phi(\mathbf{r}_1 + \mathbf{r})]^2 \rangle, \quad (2.27)$$

characterizes the differences in phase between locations  $\mathbf{r}_1$  and  $\mathbf{r}_1 + \mathbf{r}$  at distance  $r = |\mathbf{r}|$  apart in the telescope aperture. Inserting the expression for phase aberrations as a function of index of refraction from Eq. 2.26 into the structure function and separating the path length of travel from the height-only dependance of the  $C_n^2$  profile, results in the following expression for the phase structure function [42],

$$D_\phi(\mathbf{r}) = 2.91k^2(\cos\gamma)^{-1} \int C_n^2(h)dh r^{5/3}. \quad (2.28)$$

Here,  $\gamma$  is the zenith angle, the angular distance from straight up which increases the path length of light traveling through the atmosphere to enter the telescope. Fried established that the phase structure function could be expressed as a 5/3 dependance on  $|\mathbf{r}|$ ,

$$D_\phi(\mathbf{r}) = Ar^{5/3}, \quad (2.29)$$

and defined the parameter  $r_0$  as [43],

$$r_0 = \left[ 0.423k^2 \int C_n^2(h)dh \right]^{-3/5}. \quad (2.30)$$

Combining the above expressions leads to the common representation of the phase structure function,

$$D_\phi(r) = 6.88 \left( \frac{r}{r_0} \right)^{5/3}, \quad (2.31)$$

where the value 6.88 was chosen on the basis of an analysis of the performance of an optical heterodyne detection system that Fried carried out in [43]. In astronomical uses  $r_0$  is also called the coherence length and is defined to be the radius of a circle containing one radian of phase variance. Phase variance below one radian is generally considered to have low impact on image quality.

A related property is the coherence time of the atmosphere,  $\tau_0$ ,

$$\tau_0 = 0.31 \frac{r_0}{\bar{V}}, \quad (2.32)$$

where  $\bar{V}$  is a vertical average of the wind velocity over all the turbulence. This is useful for understanding how quickly an AO system must be run. The amount of temporal lag error,  $\sigma_\tau^2$  where  $\tau$  is the time lag introduced into the AO loop between measurement and command (of the DM), is given in terms of residual phase variance and can be computed for a given instrument and atmospheric coherence time, [44],

$$\sigma_\tau^2 = \left( \frac{\tau}{\tau_0} \right)^{5/3}. \quad (2.33)$$

In the case of a finite outer scale, the modified phase spatial structure function is given by [45],

$$D_\phi(r) = \left( \frac{L_o}{r_0} \right)^{5/3} \frac{2^{1/6}\Gamma(11/6)}{\pi^{8/3}} \left[ \frac{24}{5}\Gamma\left(\frac{6}{5}\right) \right]^{5/6} \left[ \frac{\Gamma(5/6)}{2^{1/6}} - \left( \frac{2\pi r}{L_0} \right)^{5/6} K_{5/6}\left(\frac{2\pi r}{L_0}\right) \right]. \quad (2.34)$$

with  $\Gamma$  the "gamma" function and  $K_{5/6}$  a modified Bessel function of the third kind. Once again, using the relationship between structure function and PSD, the spatial PSD of the phase aberrations for a given layer representing a fraction,  $f_{r_0}$ , of the total

atmosphere is given by [46],

$$W_\phi(\boldsymbol{\omega}) = \frac{\Gamma^2(11/6)}{2\pi^{11/3}} \left[ \frac{24}{5} \Gamma\left(\frac{6}{5}\right) \right]^{5/6} r_0^{-5/3} \left( \boldsymbol{\omega}^2 + \frac{1}{L_0^2} \right)^{-11/6} f_{r_0}. \quad (2.35)$$

Finally, the phase spatial covariance function for finite outer scale can be derived using Eq. 2.13,

$$\mathcal{C}_\phi(\mathbf{r}) = \left( \frac{L_0}{r_0} \right)^{5/3} \frac{\Gamma(11/6)}{2^{5/6}\pi^{8/3}} \left( \frac{2\pi\mathbf{r}}{L_0} \right)^{5/6} K_{5/6} \left( \frac{2\pi\mathbf{r}}{L_0} \right) \left( \frac{24}{5} \Gamma(6/5) \right)^{5/6}. \quad (2.36)$$

These structure functions and PSDs form the basis for both tomographic reconstruction and temporal prediction of the atmosphere based on a handful of priors and a limited number of measurement directions. The full tomographic framework is developed in detail for an MOAO system in Chapter 3.

#### 2.1.4 The effects of optical turbulence on imaging

The effects of turbulence induced focusing, spreading, scintillation, beam wander and speckles is an image that is smeared out across a detector. Consider a point source entering a telescope, the theoretical angular width of the point source on a detector is determined by the diffraction limit. A characteristic diffraction limited image of a point source has a bright inner core surrounded by dimmer Airy rings. After traveling through the atmosphere, the image of that same point source can be seen in the top left of Fig. 2.4; the photons are spread over a large area of the detector. Turning the AO system on (in this example the image is from the Canada France Hawaii Telescope's (CFHT) Hokupa'a instrument), the bright core of the point source is recovered. If the science instrument being fed by the telescope is an imager, it is obvious that the resolution of a seeing limited system is much worse than for a diffraction limited system.

If the instrument is a spectrograph, the light must pass through a narrow slit and the wider the distribution of the light, the fewer photons will pass through the slit. There are several metrics commonly used in astronomy which assess the quality of the light distribution, the first, seen in Fig. 2.4, is the Strehl ratio, (SR). This quantity is generally defined as the ratio of the peak intensity of an aberrated image of a point source compared to the maximum attainable intensity using an ideal optical system limited only by diffraction over the system's aperture. The next is Full Width at

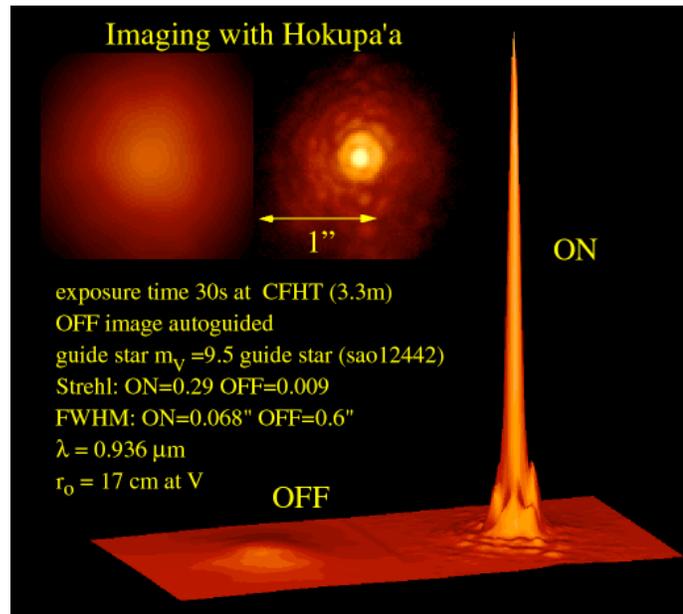


Figure 2.4: Effects of atmospheric turbulence on a star (left), AO compensated image (right)[4].

Half Max (FWHM), the width of the image, or Point Spread Function (PSF), at half of its maximum intensity. This quantifies the distribution of the light over the detector. These two quantities are typically applied to images to assess the ability of an instrument to resolve different objects. A quantity used often as the figure of merit in this dissertation is the Ensquared Energy (EE), a take-off from the perhaps more familiar quantity Encircled Energy, this is effectively the same quantity made specific to square Charged Coupled Device (CCD) pixels. The EE of an image is the ratio of the intensity within a selected region of the detector to the total intensity on the entire detector, that is, how much of the total flux has fallen within a given area. This quantity has the most meaning for spectrographs, where the shape of the PSF is slightly less critical, the goal is simply to get as much light into the slit as possible.

### 2.1.5 Angular anisoplanatism

Although the statistical properties of the atmosphere may be isotropic over a large (on the order of 10s of meters) distance and for a reasonable length of time, the instantaneous phase in two different directions is not the same. It, in fact, becomes increasingly different with distance in a way that can be approximated using the statistical tools described above, combined with site data collected during long term

surveys. This important quantity is called the isoplanatic angle; this quantity corresponds to the angle between two points on-sky at which the phase variance reaches one radian. This is a direct result of the turbulence being distributed in altitude, if all the turbulence was located at the ground, then anisoplanatism would not be a problem.

The phase variance due to anisoplanatism is written [42],

$$\sigma_{aniso}^2 \theta = 6.88 \left( \frac{\theta \bar{h}}{r_0 \cos \gamma} \right). \quad (2.37)$$

Setting this equation equal to one and solving for  $\theta$  gives the isoplanatic angle,  $\theta_0$ ,

$$\theta_0 = 6.88^{-3/5} \frac{r_0 \cos \gamma}{\bar{h}} \quad (2.38)$$

where  $\bar{h}$  is an average altitude computed from the  $C_n^2$  profile,

$$\bar{h} = \left[ \frac{\int h^{5/3} C_n^2(h) dh}{\int C_n^2(h) dh} \right]^{3/5} [43]. \quad (2.39)$$

When two objects are separated by an angle greater than the isoplanatic angle, the measured WFs from one become sufficiently different from the other that using the measurement from the first as a direct estimate of the WF of the other does not provide good AO correction, they are suffering the effects of anisoplanatism (see Fig. 2.6 below). This is the fundamental limitation of classical SCAO systems which both limits their sky coverage and prevents them from providing wide field correction and thus is the main motivation for developing higher complexity specialized WFAO instruments.

### 2.1.6 Laser Guide Stars

One method developed to increase the sky coverage of AO systems is to create artificial reference sources that can be pointed toward a science object within its isoplanatic patch. This is done with powerful laser beacons that excite specific atoms or molecules in the atmosphere, causing them to re-emit the light thus generating an artificial star within the atmosphere. There are two types of astronomical LGSs, Rayleigh guide stars typically use green wavelengths and excite molecules at altitudes around 15km. Sodium LGSs use the specific 589nm yellow wavelength to excite sodium atoms that exist in a layer surrounding the earth at an altitude of approximately 90km. LGSs

are in use at many major observatories to improve the sky coverage of their SCAO systems; they are also beginning to be implemented not just as single lasers but as constellations for WFAO instruments. Several of the facility AO systems planned for ELTs are designed to use constellations of LGSs to do MCAO. The Gemini MCAO System (GeMS) has successfully carried out MCAO using a constellation of sodium LGSs [47].

There are limitations to using LGSs which prevent full sky coverage, most notably, they cannot provide information on atmospheric tip/tilt (T/T), the lowest order spatial modes which cause global motion of an object (wander), due to the nearly instantaneous travel of the laser light up then back down through the same volume of turbulence. LGS systems thus require supplementary information from an NGS to acquire the T/T measurements. These T/T NGSs do not need to be as bright, nor as close to a science object to be effective and as such the sky coverage is still greatly increased. Another limitation is the cone effect; the LGS is generated at a finite altitude within the atmosphere and enters the telescope as an expanding cone of light, whereas light from astronomical objects effectively originates at infinity and passes through a cylinder of atmospheric turbulence, entering the telescope as a collimated beam. Some of the turbulence affecting the astronomical light will not be sensed by the LGS light, resulting in an incomplete estimate of the phase aberrations.

Subaru telescope can generate an on-axis sodium LGS, and it is intended to try to use it in concert with NGSs, either with three NGSs to improve performance or with two NGSs to increase the sky-coverage of the Raven system.

### 2.1.7 Representing the phase

The phase aberrations must be represented mathematically in a chosen basis. This dissertation will present work carried out in both zonal and modal bases. The fundamental tools of the modal basis representation are presented here.

Zernike polynomials are an orthogonal set of polynomials defined on a unit circle; a phase aberration can be represented by a weighted sum of these polynomials. They are a very useful and intuitive method for describing the aberrations of an optical system and can be visualized as various shapes as shown in Fig. 2.5. Noll [6] described a set of Zernike polynomials which is now typical to AO applications:

$$\left. \begin{aligned} Z_{evenj} &= \sqrt{n+1} R_n^m(r) \sqrt{2} \cos m\theta \\ Z_{oddj} &= \sqrt{n+1} R_n^m(r) \sqrt{2} \sin m\theta \end{aligned} \right\} m \neq 0 \quad (2.40)$$

$$Z_j = \sqrt{n+1}R_n^0(r), \quad m = 0$$

where

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s![(n+m)/2-s]![(n-m)/2-s]!} r^{n-2s}, \quad (2.41)$$

and  $n, m \in \mathbb{Z}$  are respectively the radial and azimuthal order of the mode and satisfy  $m \leq n$ ,  $n - |m| = \text{even}$ . The index  $j$  is a function of  $m$  and  $n$ . Noll also gives the expression for the Fourier transform of Zernike polynomials,

$$Q_j(r, \phi) = \sqrt{n+1} \frac{J_{n+1}(\pi f D)}{\pi f D} \begin{cases} (-1)^{(n-m)/2} i^m \sqrt{2} \cos m\phi, & (\text{even } j) \\ (-1)^{(n-m)/2} i^m \sqrt{2} \sin m\phi, & (\text{odd } j) \\ (-1)^{n/2}, & (m = 0), \end{cases} \quad (2.42)$$

a set of functions that will be useful in the analysis of both the spatial and temporal behaviour of the atmosphere.

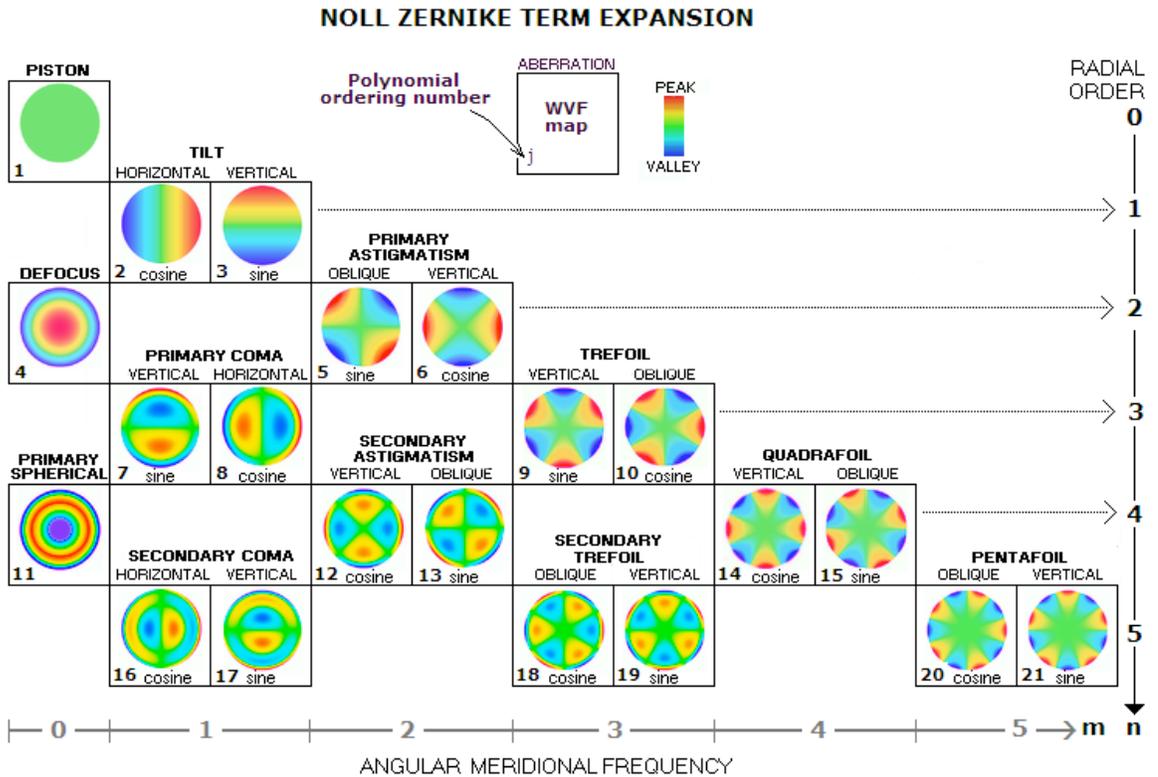


Figure 2.5: Visualization [5] of the first few radial orders of Zernike polynomials in the ordering given by Noll [6].

## 2.2 Adaptive Optics technologies

A closed-loop SCAO system encompasses most of the key technology and control elements required by any AO system. A typical closed loop AO system layout, shown in Fig. 2.6, shows a WFS, a DM and a feedback path with a controller. Also illustrated is the angular separation between science object and GS, leading to error from anisoplanatism as previously described in Sec. 2.1.5.

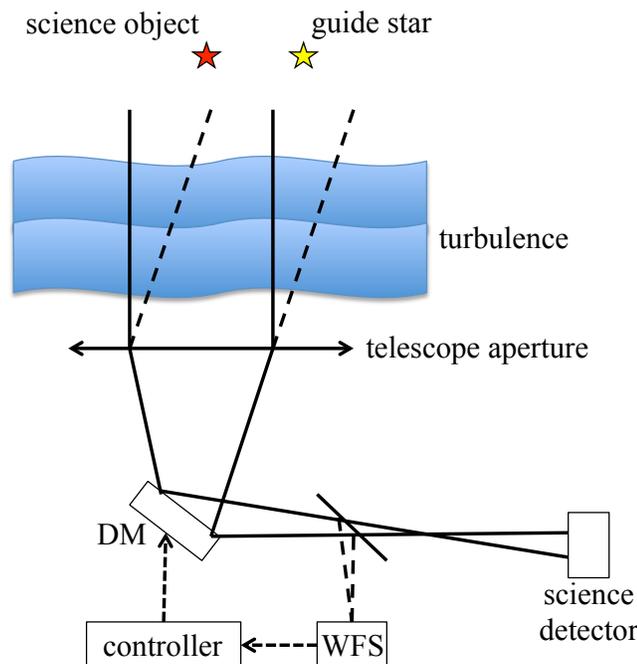


Figure 2.6: Basic closed loop schematic for a single conjugate AO system.

The role of the WFS is to collect the incident light from the reference source in a way which enables the WF phase aberrations to be estimated (reconstructed) by relating it to intensity. A clear relationship between phase aberrations and WFS measurements must therefore exist and be known. There are multiple categories of WFS devices used in AO, some notable ones are: Curvature sensors, pyramid WFSs, and Shack-Hartmann WFSs (SH-WFS). The curvature WFS works by comparing the illumination patterns as a function of the two-dimensional image-plane co-ordinate,  $\mathbf{r}$ , in a pair of images,  $I_1(\mathbf{r}), I_2(\mathbf{r})$ , one located before the focal plane (intra-focal) at distance  $l$  and one located after the focal plane (extra-focal) at the same distance. The pyramid WFS is a four-sided pyramid which splits the image into four sub-images. From these sub-images, the local gradient of the phase can be determined

by comparing the intensity on the pixel corresponding to the same location in the aperture in each sub image. The WFS which was used throughout this work is the SH-WFS; this choice was made for a variety of reasons, mainly driven by the maturity of the technology, the limitations of other devices and the collective experience of the Raven team members in working with the SH-WFS.

### 2.2.1 The Shack-Hartmann WFS

The SH-WFS is the device used for this work, both in simulation and on the instrument. It consists of a grid of small lenses termed lenslet array. Each small lens focuses a portion of the incident wavefront onto a CCD detector, resulting in a grid of spots. If the incident wavefront is flat, this grid is regular; any aberrations in the wavefront will cause the spots to move relative to their reference position by an amount proportional to the angle of arrival of the portion of wavefront focused by each lenslet, as shown in Fig. 2.7.

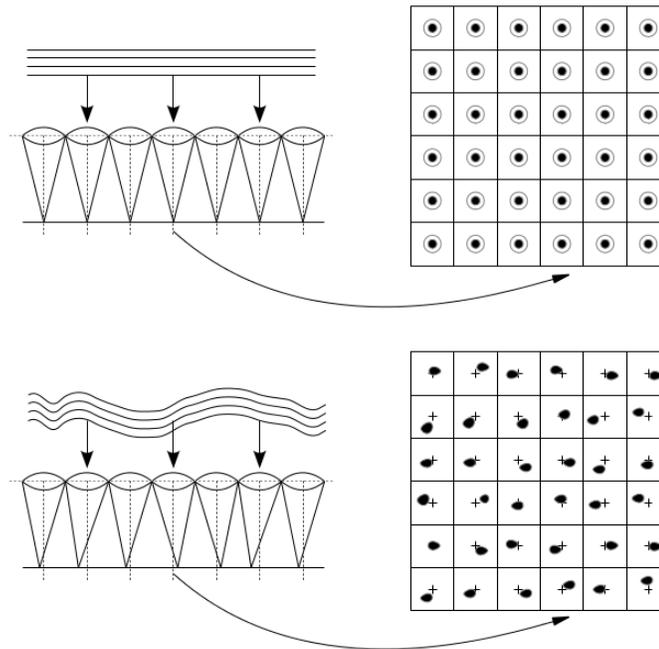


Figure 2.7: The effects of turbulence on SH-WFS spot positions. Plane waves (top) focus to diffraction limit at the centre of the lenslets; aberrated wavefronts (bottom) cause distortion in spot shape and overall displacement of the spot centres.

There are many techniques, called centroiding algorithms, to determine the position of each spot on its corresponding subaperture of the detector. For example, the

Centre of Gravity (CoG) centroiding algorithm computes the location of a spot in a subaperture that is  $n_p$  by  $n_p$  pixels is given by,

$$c_x = \frac{\sum_{i,j} x_{i,j} I_{i,j}}{\sum_{i,j} I_{i,j}}; \quad c_y = \frac{\sum_{i,j} y_{i,j} I_{i,j}}{\sum_{i,j} I_{i,j}}. \quad (2.43)$$

It is a function of the position of the pixels within the lenslet,  $(x_{i,j}, y_{i,j})$ , and the corresponding intensity of the pixel,  $I_{i,j}$ . Sources of noise include photon noise, proportional to  $\sqrt{N}$  where  $N$  is the number of photons, and read-out noise due to conversion of electron detection events to voltage readings; these reduce the accuracy of the CoG computation. First order solutions to reduce noise errors include windowing, in which a fixed number of pixels surrounding the spot location on the sub-aperture are selected and used for centroiding. Thresholding is also an effective technique where a level, based on the known noise properties of the system, can be specified and all pixel intensities falling below that level are set to zero. More sophisticated techniques to deal with noise as well as spot distortion or elongation include correlation centroiding [48] and matched filtering [49]. WFS spot elongation will occur be significant in the case of LGS spots on ELTs; it also occurs for 8m class telescopes such as Subaru, but is small and will not be treated as elongation on Raven.

### 2.2.2 Deformable mirrors

DM technology typically consists of a flexible mirror membrane located over a grid of controllable micro-actuators. There are several types of DM actuation technology: piezo-electric actuators, bimorph membranes, voice coil actuators and MEMS. There are pros and cons to all of these technologies and the trade-off between actuator stroke, bandwidth, actuator pitch and number of actuators is a balance that must be evaluated depending on the requirements of the instrument. In general, small pitch, high bandwidth and high order mirrors (such as a MEMS) will not have as much stroke, or as large an aperture as a lower order DM with a lower bandwidth. In some cases, high and low order DMs have been combined in a woofer-tweeter configuration [50] in order to better correct the entire spectrum of turbulence. This is possible because the low order atmospheric aberrations also happen to be the largest (requiring the most stroke), and evolve more slowly in time compared to the higher order aberrations.

The major contribution made by the DM to the overall system error budget is

fitting error,  $\sigma_f^2$ . This error is again given in terms of residual phase variance and can be computed for a given DM-instrument pair [44],

$$\sigma_f^2 = k \left( \frac{D}{r_0} \right)^{5/3} N^{-5/6} \quad (2.44)$$

$$(2.45)$$

The constant  $k$  is a factor dependant on the coupling of the DM influence functions,  $D$  is the pupil diameter, and  $N$  is the total number of actuators. The total amount of stroke (in  $\mu m$ ) required to compensate for the atmosphere is given by,

$$\delta = \frac{3\lambda}{2\pi} \sqrt{l} \left( \frac{D}{r_0} \right)^{5/6}. \quad (2.46)$$

In this expression,  $l$  can have a value of 1.03 if the DM is meant to correct atmospheric T/T, or 0.134 if a separate T/T correction device is included in the instrument. To avoid actuator saturation in bad seeing conditions, a worst-case scenario  $r_0$  value is used. In addition, some stroke must be added to accommodate correction of wind-shake, as well as static and quasi-static errors in the system such as alignment errors, gravitational flexure, DM flattening and drift.

### 2.2.3 Command estimation for SCAO

The specific problem in SCAO is to compute DM commands,  $\mathbf{u}$ , from a sampled map of local gradients or other measurements,  $\mathbf{s}$  via a measured linear relation. The basic static linear measurement model,

$$\mathbf{s} = \mathbf{IM}\mathbf{u} + \eta, \quad (2.47)$$

indicates that this computation is an inverse problem. Throughout this dissertation, vectors are indicated by bold lower case characters and matrices by bold upper case characters. The DM actuator command vector has length  $m_u$  and is typically smaller than the measurement vector of length,  $n_s$ . This means that the matrix  $\mathbf{IM}$  which relates the two quantities is of dimensions  $n_s \times m_u$  where  $n_s > m_u$ .  $\mathbf{IM}$  is called the interaction matrix and can be measured directly from the system by pushing each DM actuator sequentially (or in a more complex arrangement) and recording that actuator's influence on the WFS (measuring the slopes). Typically, bright calibration

sources are used to take these measurements and we can therefore neglect any measurement noise. One can then invert the interaction matrix to obtain the command matrix. Because  $\mathbf{IM}$  is not square, this is done via its singular value decomposition (SVD),

$$(\mathbf{IM}) = \mathbf{USV}^T, \quad (2.48)$$

$$(\mathbf{IM})^\dagger = \mathbf{VS}^{-1}\mathbf{U}^T, \quad (2.49)$$

where  $\mathbf{S}$  is a diagonal matrix containing the singular values of  $\mathbf{IM}$  and can therefore be inverted by taking the inverse of the diagonal values. Thresholding can be carried out at this step to prevent inverting singular values that are very small which would lead to large values in  $\mathbf{S}^{-1}$ . All singular values below a certain threshold can be set to zero, this avoids the amplification of poorly sensed actuator influence, especially near the edge of the pupil.  $\mathbf{V}$  and  $\mathbf{U}$  are non-square matrices with dimensions  $m_u \times n_e$  and  $n_s \times n_e$  respectively with  $n_e$  the number of eigenvalues maintained by  $\mathbf{S}$  after thresholding.

## 2.2.4 Principles of WF estimation for WFAO

Under the hypothesis that the turbulent atmosphere is a sum of  $N_l$  thin layers located at a discrete number of different altitudes  $h_l$ , the aperture-plane phase  $\phi(\boldsymbol{\rho}, \boldsymbol{\theta}, t)$  indexed by the bi-dimensional spatial coordinate vector  $\boldsymbol{\rho} = (\rho_x, \rho_y)$  in direction  $\boldsymbol{\theta} = (\theta_x, \theta_y)$  at time  $t$  is defined as

$$\phi(\boldsymbol{\rho}, \boldsymbol{\theta}, t) = \sum_{l=1}^{N_l} \mathcal{W}_l(\boldsymbol{\rho} + h_l \boldsymbol{\theta}, t) \quad (2.50)$$

where  $\mathcal{W}_l(\boldsymbol{\rho}, t)$  is the  $l^{\text{th}}$ -layer wave-front. The aperture-plane phase is not measured directly in most AO systems and the WF phase is reconstructed from a set of discrete measurements using a measurement model. The SH-WFS introduced in Sec. 2.2.1, which provides phase gradients,  $\mathbf{s}$  with measurement noise, will be used from now on.

MOAO systems require an open-loop estimate of the atmosphere over a large field in a discrete number of correction directions based on WFS data from multiple measurement directions;  $\mathbf{s}_\alpha$  are noisy measurements made in specific GS directions,

$$\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1 \dots \boldsymbol{\alpha}_n]. \quad (2.51)$$

When developing the reconstructor in Chapter 3, we will first consider the static case, which indicates that the solution relates only to the latest available measurement, *i.e.* the temporal evolution is not seized by a static algorithm. WF reconstruction is carried out in the pupil-plane and temporal aspects of the system are ignored, therefore the simplified SH-WFS measurement model can be used,

$$\mathbf{s}_\alpha = \mathbf{\Gamma}\phi_\alpha + \boldsymbol{\eta} \quad (2.52)$$

with  $\mathbf{s}_\alpha \in \mathfrak{R}^{(n_\alpha \times n_s) \times 1}$  a column-vector of  $n_\alpha \times n_s$  measurements for all the  $n_\alpha$  directions, obtained using the pupil-plane gradient operator  $\mathbf{\Gamma} \triangleq \text{diag}\{[\Gamma_1, \dots, \Gamma_n]\} \in \mathfrak{R}^{(n_\alpha \times n_s) \times (n_\alpha \times N_z)}$  that concatenates  $n_\alpha$  individual pupil-plane  $\Gamma_{\alpha_i}$  relating the  $N_z$  phase dimensions in the aperture in the directions  $\boldsymbol{\alpha}$  to the  $n_s$  measurements. Noise is represented by  $\boldsymbol{\eta}(t) \triangleq \in \mathfrak{R}^{(n_\alpha \times n_s) \times 1}$  and is assumed to be a gaussian zero-mean white noise process.

## 2.3 Raven: A Multi-Object Adaptive Optics test bed

As stated in the introduction, Raven will be the first MOAO instrument on an 8 m class telescope feeding an AO- optimized science instrument. It is a science and technology pathfinding instrument, whose purpose is to establish the feasibility of proposed solutions to known challenges, and to identify and solve any new or unexpected problems, thus reducing the overall risk of constructing a facility MOAO instrument.

Much effort was initially invested in establishing the baseline parameters of Raven [51, 52] and the opto-mechanical design [53]. A functional block diagram of Raven's optical system is shown in Fig. 2.8.

Raven consists of 8 main optical subsystems:

- The deployable Calibration Unit (CU) is a telescope simulator and a turbulence generator. It also contains an array of off-axis sources and one on-axis LGS source. The intended functions of the CU are to: 1) help align other Raven subsystems, 2) calibrate the AO system (generate interaction matrices and measure field-dependent non-common path aberrations), and 3) test the MOAO system by including three phase screens, one of which is a  $17 \times 17$

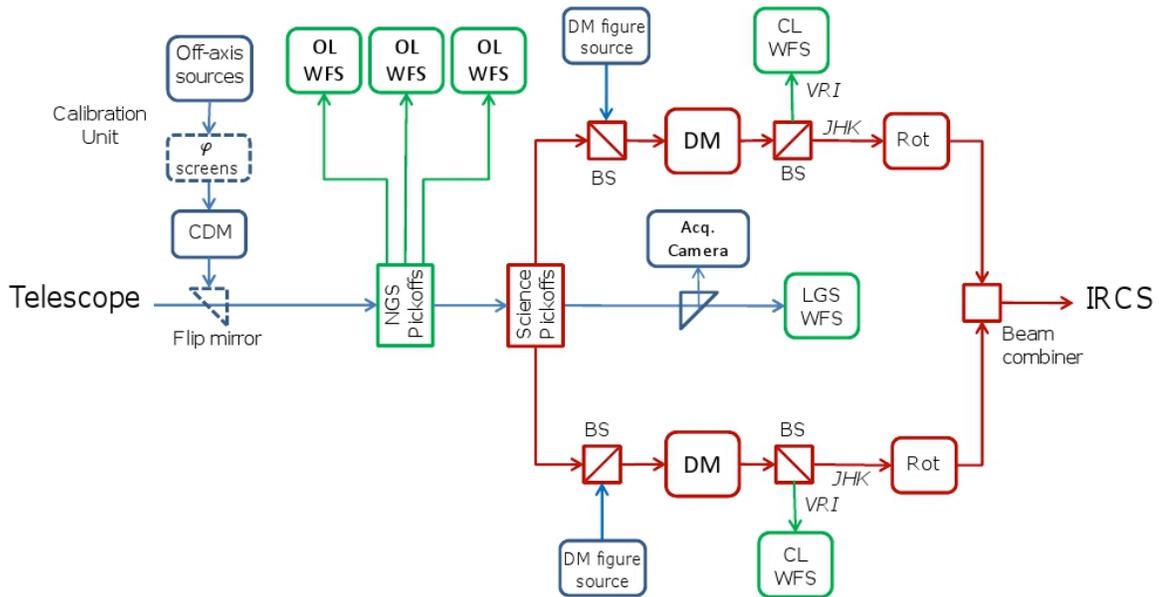


Figure 2.8: Functional optical block diagram of RAVEN. Dashed blocks are deployable. Raven consists of 8 main subsystems: the deployable Calibration Unit, the Open-Loop NGS WFSs, the Science Pick-offs, the Science Relays, the Closed-Loop NGS Truth/Figure WFSs, the Beam Combiner, the LGS WFS and the Acquisition Camera.

ground-conjugated DM called the Calibration DM (CDM), which by design can generate turbulence above the (spatial) Nyquist frequency of the WFSs. Laboratory experiments referred to in later sections will be carried out using the CU as a telescope simulator.

- Three NGS OL-WFSs are mounted on deployable pickoff arms. These are mounted on x-y translating stages to manoeuvre them into the positions of the NGSs within the field of regard as well as track the stars as the field rotates. It should be noted that this field rotation is due to the inability of Raven to use the Subaru image de-rotator as a result of space constraints on the Naysmith platform. The rotation is not a fixed configuration rotation, that is, the axes of the WFSs remain fixed and their positions relative to each other change. This has ramifications on the on-sky tomography which requires current knowledge of the relative positions of the WFSs.
- On-axis LGS WFS will be fed by the Subaru Sodium beacon in order to improve AO correction and/or the sky coverage.

- Two science pick-off arms whose design consists of a mirror mounted on an  $r-\theta$  arm followed by a trombone mirror that keeps the path length constant. The science relay for each arm contains a DM which is a custom ALPAO DM with 13x13 actuators and a 25 mm aperture.
- A figure source and closed-loop (CL)-WFS share the science relay optical path in each arm and can be used to either: 1) measure the shape of the DM using the figure source; this can be done in real time at a rate equal to or possibly greater than the system frame rate, 2) use the CL WFS as a truth WFS to measure the MOAO performance or contribute low temporal frequency correction information from longer exposures if the science object is compact, or 3) use the CL WFS as a classical AO system that uses the science target as the NGS.
- A K-mirror in each science relay which can rotate the images of the science targets so that extended objects can be properly aligned onto the slit of IRCS.
- The Roof mirror combines light from both arms of the system so that the common beam shares an identical exit pupil and provides two adjacent 4 arc-second science fields to the single IRCS slit.
- An acquisition camera is used to determine the telescope pointing and facilitate target acquisition by ensuring that shadows of the probe arms fall over the NGSs and science targets.

Some of the main technical specifications are provided in Table 2.1.

## 2.4 Wavefront reconstruction modes on Raven

MOAO has the potential to deliver near diffraction-limited images to multiple, small patches spread across a large FoR. One challenge of an MOAO system is that it is highly distributed; for Raven, light from up to three NGSs and one LGS will be sensed within a 3.5 arc-min FoR. Pixels from the OL-WFS detectors will be read by the RTC and transformed into a tomographic model of the atmosphere above the observatory. This tomographic model will be sampled in directions defined by the position of the science probes in the patrol field and DM commands will be generated and applied. All of these actions are performed using OL control. Accurate knowledge

Table 2.1: Raven specifications.

<b>Calibration Unit</b>	
Telescope D	8 m
$r_0$	15.6 cm
$L_0$	40 m
Fractional $r_0$	[0.596; 0.224; 0.180]
Altitudes	[0, 5, 10] km
wind speeds	[5.68; 6; 17] m/s
wind direction	[90; 180;180] deg
<b>Wavefront Sensor</b>	
RON	$0.2 e^-$
$N_{NGS}$	3
$N_{LGS}$	1
Order	$10 \times 10$
$\theta_{pix}$	0.4 arcsec
$N_{pix}$	12
framerate	up to 500Hz
<b>DMs</b>	
Science DMs	2
Order	$13 \times 13$
Number of valid actuators	145
Calibration DM	1
Order	$17 \times 17$
Number of valid actuators	225

of the science probe placement in the focal plane and the relative alignment of the DMs and WFSs in the pupil plane is required.

The primary research objective has been to bring a new perspective to tomographic wavefront reconstruction for MOAO while giving consideration to the future directions of the technology. This has required attention to computational complexity and provided proof, both in simulation and in practice, that good performance can be achieved for multiple science channels acting in parallel. There are several operational modes used on Raven which have allowed direct comparison of various OL wavefront reconstruction techniques.

### 2.4.1 Closed Loop Mode

Raven has a CL WFS in each of the science pickoff arms. If the science target(s) are bright stars, Raven can operate as a classical AO system with on-axis correction. As the best AO correction Raven can possibly achieve, the closed-loop performance will be used as a reference for the open-loop performance.

### 2.4.2 Ground Layer AO

In order to show that the added complexity of doing tomography with MOAO is advantageous, we must show definitively that it improves performance compared to GLAO. The GLAO mode is easily implemented within the RTC by bypassing the tomographic reconstructor completely and implementing an averaging of the OL-WFS measurements.

### 2.4.3 Static MOAO Mode

This is the primary MOAO validation tool, it uses a static tomographic reconstructor and has been used extensively in the lab to characterize system performance and prove that Raven was ready for on-sky testing. In this mode, the OL-WFSs are deployed and feed measurements to the RTC. The science pickoff arms are also deployed and feed light to IRCS (or the science camera). The RTC reads the OL-WFS pixels, measures the current slopes, and performs a tomographic reconstruction based on the location of the OL-WFS probe arms and the science pickoffs. DM commands are generated and sent to the devices.

### 2.4.4 MOAO with CL-WFS Mode

The Raven opto-mechanical design includes a CL-WFS on each science path (Fig. 2.9) to facilitate calibration [54, 55] and provide a baseline; however, we can take advantage of these design components which may also be used during observations to provide supplementary information on quasi-static errors.

In certain cases, the science target itself may be bright and compact enough to take slow CL-WFS measurements of the phase in the science direction. This is not enough to do closed loop AO, but the information can be used to augment the accuracy of the wavefront estimation. It is an important research objective to design the data

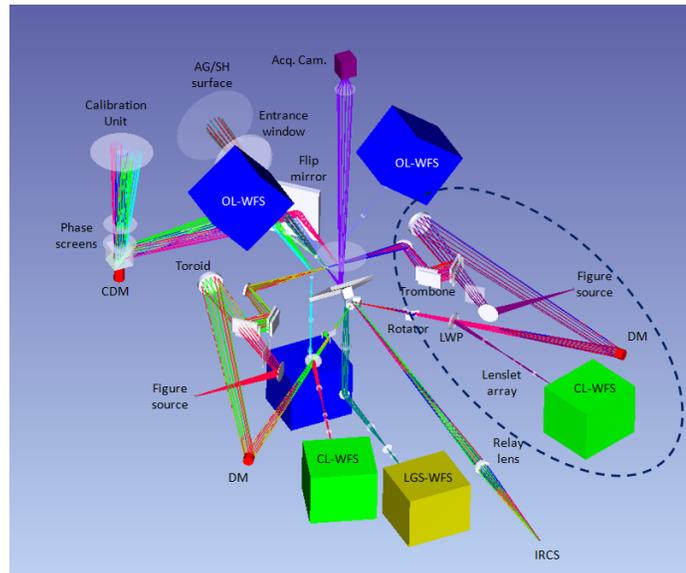


Figure 2.9: 3D representation of Raven. Highlight on one CL-WFS path and figure source which will enable more WF correction modes.

processing pipeline and control system to incorporate this information in an optimal way as it becomes available.

#### 2.4.5 MOAO with Figure WFS mode

The system also includes a reference source on each science path. The same static MOAO reconstruction process as in Sec. 2.4.3 is carried out, and in addition, the reference source, or figure source illuminates the DM and CL-WFS only and thus provides a measurement of the DM shape. The figure source is a bright, visible wavelength point source. It is independent of the science target and pick-off arm and can therefore be run very fast (up to 500Hz). The wavefront reconstruction will generate a set of DM commands which will represent an expected DM shape according to the calibrated closed loop interaction matrix; the Figure WFS will be able to measure the actual DM shape. Comparing the measured shape to the expected shape, a local closed loop can be used to drive the residual to zero. This will reduce risk posed to open loop systems by DM go-to error and calibration errors. At the very least, the slopes from the Figure WFS can be archived to compare to what is expected in post-processing.

Note that all of the AO modes mentioned above except for CL Mode can be executed either with 3 NGSs, 2 NGSs plus the LGS or all 3 NGSs and the LGS.

## Chapter 3

# Tomography for MOAO

### 3.1 Sky coverage

MOAO systems achieve their scientific potential by taking advantage of their multiplexing observing capabilities. The availability of sufficiently bright guide-stars overshadows the utility of AO, even more so when the main scientific goals involve resolving photometric and kinematic science observations concentrating the light sufficiently to obtain spectra for several targets in a reasonable amount of time. For a more in-depth description of science cases for MOAO systems see [56]. Since such targets are usually very faint and not suitable for guiding the AO system, other natural guide-stars must be found that are sufficiently close. Therefore, the sky-coverage, *i.e.* the percentage of available sky for observing, imposes a strong constraint on the observable targets.

As a pathfinder instrument intended to perform scientific observations, it is of key importance that interesting science targets are available to Raven. Employing a static tomographic reconstructor, Raven's limiting magnitude will approach 14.5 (for 30% EE) using a reduced frame rate of 180Hz. As stated in [51], sky coverage for Raven will be low. As an example, consider a point with galactic coordinates  $(b; l) = (30; 0)$ . Using the Besançon model of the galaxy [57], one finds that there are 750 stars per square degree with  $R < 14.5$  (1040 stars per square degree with  $R < 15$ ). The probability that there are three stars with  $R < 14.5$  in a 2 arc-min diameter FoR is just 3%. This does not even account for asterisms that are unsuitable for Raven; in some cases the science targets will not be inscribed within the potential NGS asterism and, therefore, the tomographic error will be too great. Star densities are increasing

as a power law at these magnitudes, so going 1 magnitude deeper can increase the density of available stars by 1.8.

There are several ways to approach the challenge of increasing the limiting magnitude, these include longer integration periods on the WFSs to increase SNR, better centroiding algorithms to obtain more accurate slope measurements in low signal regimes, and better noise rejection in the tomographic reconstructor itself. Ideally, an optimal combination of these ideas can be achieved. The development of tomographic reconstructors which can operate effectively in the lower WFS framerate regime, employ better noise rejection and the resulting potential improvement in terms of Raven’s limiting magnitude are the subject of this chapter. More sophisticated centroiding algorithms can (and will) be used in conjunction with the methods developed here, but their development and analysis is outside the scope of this work.

This chapter expands and extends the material published in [58] (Appendix B). It starts by presenting an established solution to the AO minimization criterion in the context of MOAO and an examination of the information made available by this solution under different computational approaches is presented. By increasing the WFS integration times to increase SNR, a tradeoff occurs whereby temporal lag error is increased, potentially off-setting any gain accrued by improving the SNR. To counteract this trade-off, a temporal prediction step can be carried out in an effort to reduce the effect of the increased lag. Several methods of temporal prediction in both modal (Zernike) and zonal bases are developed and compared. The temporal prediction model can also be combined with a Kalman filter to better handle spatial errors, due to noisy slopes, and to further improve performance with high magnitude GSs.

## 3.2 Static Minimum Mean Square Error

In MOAO, in order to maximize SR and EE, the objective cost function is the minimization of the aperture-plane residual phase variance in selected science directions,  $\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \dots \boldsymbol{\beta}_{N_\beta}]$ ; an estimate is made for each direction,  $\boldsymbol{\beta}_i$ ,

$$\mathbf{R} = \arg \min_{\mathbf{R}'} \left\langle \|\boldsymbol{\phi}(\boldsymbol{\beta}_i) - \hat{\boldsymbol{\phi}}(\boldsymbol{\beta}_i)\|_{L_2(\Omega)}^2 \right\rangle \quad (3.1)$$

where  $\phi$  is the actual phase (represented in an arbitrary basis),  $\hat{\phi}(\beta_i) \triangleq \mathbf{R}\mathbf{s}_\alpha$ , is the estimated phase (estimated quantities are indicated by the hat symbol),  $L_2$  is the Euclidean norm over the aperture  $\Omega$  and  $\langle \dots \rangle$  is the ensemble average over time for an individual optimization direction.  $\mathbf{s}_\alpha$  are noisy measurements made in specific GS directions,

$$\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1 \dots \boldsymbol{\alpha}_n]. \quad (3.2)$$

Recall that integrated phase in the telescope pupil plane will be represented by  $\phi$  and the distributed phase in a discrete number of atmospheric layers will be expressed as  $\varphi$  which, in the case of a modal basis, is a concatenated vector of Zernike coefficients representing the phase, and in the case of a zonal basis is a concatenated vector of phase values.

A well established solution to the minimization criterion in Eq. 3.1 is the Minimum Mean Square Error (MMSE) solution [21, 59, 60, 61]. It states that for two jointly Gaussian, zero mean, random variables (in this case  $\mathbf{s}_\alpha$  and  $\phi_\beta$ ) with covariances,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{\phi_\beta, \phi_\beta} & \boldsymbol{\Sigma}_{\phi_\beta, \mathbf{s}_\alpha} \\ \boldsymbol{\Sigma}_{\mathbf{s}_\alpha, \phi_\beta} & \boldsymbol{\Sigma}_{\mathbf{s}_\alpha, \mathbf{s}_\alpha} \end{bmatrix}, \quad (3.3)$$

the estimate,  $\hat{\phi}_\beta$ , of the value taken by  $\phi_\beta$  for a given  $\mathbf{s}_\alpha$  which minimizes the Mean Square Error of Eq. 3.1 can be written,

$$\hat{\phi}_\beta = \boldsymbol{\Sigma}_{(\phi_\beta, \mathbf{s}_\alpha)} \boldsymbol{\Sigma}_{\mathbf{s}_\alpha}^{-1} \mathbf{s}_\alpha \quad (3.4)$$

where, in general,  $\beta \neq \alpha$  for practical reasons, and where  $\hat{\phi}_\beta$  is the pupil-plane phase estimate representing the decomposition of the WF in all the m- $\beta$  science directions. In practice, the measurements are converted to phase space, and the covariance matrices in Eq. 3.3 are computed according to the expression given in Eq. 2.36 for all baselines,  $|\mathbf{r}|$  between the phase points.

### 3.2.1 Explicit layered static MMSE

In many tomographic AO systems the forward measurement model makes explicit use of the layered phase vector (be it phase points or modal coefficients [59], [15]). Using matrix formulation, the resulting aperture-plane WF  $\phi(\boldsymbol{\theta}, t)$  in the near-field approximation relates to the representation of the WF phase  $\varphi$  defined over a discrete

number of layers in the volume by a simple matrix multiplication [15],

$$\phi(\boldsymbol{\theta}, t) = \mathbf{P}_{\boldsymbol{\theta}}\boldsymbol{\varphi}(t) \quad (3.5)$$

where  $\mathbf{P}_{\boldsymbol{\theta}}$  is a propagation matrix which crops the phase in each layer in direction  $\boldsymbol{\theta}$  and makes the sum of all the layers and,

$$\boldsymbol{\varphi} \triangleq \left[ \varphi_0 \quad \cdots \quad \varphi_{h_L} \right]^T. \quad (3.6)$$

This is a static method, therefore time independent; the measurement model can now be written,

$$\mathbf{s}_{\alpha} = \mathbf{\Gamma}\mathbf{P}_{\alpha}\boldsymbol{\varphi} + \boldsymbol{\eta}. \quad (3.7)$$

The MMSE reconstructor providing the aperture-plane phase estimate in the  $\boldsymbol{\beta}$  science directions via the explicit estimation of phase in the layers is

$$\hat{\boldsymbol{\phi}}_{\boldsymbol{\beta}} = \mathbf{P}_{\boldsymbol{\beta}}\mathbf{R}\mathbf{s}_{\alpha}. \quad (3.8)$$

By using the phase in each layer in the forward model given in Eq. 3.7, the estimation breaks into two steps: a 3D tomographic estimation followed by a linear propagation into the science directions.

The derivation of the explicit static MMSE reconstructor can be made using terms that can be measured or computed analytically by developing the terms in Eq. 3.4. Taking the definition of covariance [61],

$$\boldsymbol{\Sigma}_{x,y} = \langle \mathbf{x}\mathbf{y}^T \rangle, \quad (3.9)$$

and defining the covariance matrices in terms of the measurement model in Eq. 3.7 and the phase estimate in Eq. 3.8 yields the following terms,

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{s}_{\alpha}} &= \langle \mathbf{s}_{\alpha}\mathbf{s}_{\alpha}^T \rangle \\ &= \langle (\mathbf{\Gamma}\mathbf{P}_{\alpha}\boldsymbol{\varphi} + \boldsymbol{\eta})(\mathbf{\Gamma}\mathbf{P}_{\alpha}\boldsymbol{\varphi} + \boldsymbol{\eta})^T \rangle \\ &= \mathbf{\Gamma}\mathbf{P}_{\alpha} \langle \boldsymbol{\varphi}\boldsymbol{\varphi}^T \rangle \mathbf{P}_{\alpha}^T \mathbf{\Gamma}^T + \langle \boldsymbol{\eta}\boldsymbol{\eta}^T \rangle \end{aligned} \quad (3.10)$$

$$= \mathbf{\Gamma}\mathbf{P}_{\alpha}\boldsymbol{\Sigma}_{\boldsymbol{\varphi}}\mathbf{P}_{\alpha}^T \mathbf{\Gamma}^T + \boldsymbol{\Sigma}_{\boldsymbol{\eta}} \quad (3.11)$$

and

$$\begin{aligned}
\Sigma_{(\phi_\beta, s_\alpha)} &= \langle \phi_\beta s_\alpha^\top \rangle \\
&= \langle \mathbf{P}_\beta \boldsymbol{\varphi} (\mathbf{\Gamma} \mathbf{P}_\alpha \boldsymbol{\varphi} + \boldsymbol{\eta})^\top \rangle \\
&= \mathbf{P}_\beta \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle \mathbf{P}_\alpha^\top \mathbf{\Gamma}^\top + \mathbf{P}_\beta \langle \boldsymbol{\varphi} \boldsymbol{\eta}^\top \rangle \\
&= \mathbf{P}_\beta \Sigma_\varphi \mathbf{P}_\alpha^\top \mathbf{\Gamma}^\top.
\end{aligned} \tag{3.12}$$

It is assumed that no statistical correlation exists between the phase vector and the measurement noise, therefore  $\langle \boldsymbol{\varphi} \boldsymbol{\eta}^\top \rangle = 0$ . Combining the two matrices, the expression for the pupil-plane phase estimate in the science directions is written,

$$\hat{\phi}_\beta = \mathbf{P}_\beta \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle \mathbf{P}_\alpha^\top \mathbf{\Gamma}^\top (\mathbf{\Gamma} \mathbf{P}_\alpha \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle \mathbf{P}_\alpha^\top \mathbf{\Gamma}^\top + \langle \boldsymbol{\eta} \boldsymbol{\eta}^\top \rangle)^{-1} s_\alpha, \tag{3.13}$$

noting that the simplified notation for the direction-independent phase covariance in the layers,  $\Sigma_\varphi = \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle$ , and for the noise covariance,  $\Sigma_\eta = \langle \boldsymbol{\eta} \boldsymbol{\eta}^\top \rangle$ , will be used in later sections.

### 3.2.2 Spatio-Angular static MMSE

The term spatio-angular (SA) reconstructor, was coined by Rodolphe Conan, to the best of the author's knowledge, on account of the nature of the covariance matrices involved in its definition. It can be seen as a generalization of the work of Whitely et al [62], in seeking the optimal anisoplanatic reconstructor in classical AO, to the tomographic, multiple sensor case. Using the SH-WFS measurement model from Eq. 2.52, the pupil-plane estimate of the phase in the science directions can be made in terms of the slopes measured in the GS directions: The components of the MMSE solution,  $\mathbf{R}$  are computed as

$$\begin{aligned}
\Sigma_{s_\alpha} &= \langle s_\alpha s_\alpha^\top \rangle \\
&= \langle (\mathbf{\Gamma} \phi_\alpha + \boldsymbol{\eta}) (\mathbf{\Gamma} \phi_\alpha + \boldsymbol{\eta})^\top \rangle \\
&= \mathbf{\Gamma} \langle \phi_\alpha \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top + \langle \boldsymbol{\eta} \boldsymbol{\eta}^\top \rangle
\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}
\Sigma_{(\phi_\beta, \mathbf{s}_\alpha)} &= \langle \phi_\beta \mathbf{s}_\alpha^\top \rangle \\
&= \langle \phi_\beta (\mathbf{\Gamma} \phi_\alpha + \boldsymbol{\eta})^\top \rangle \\
&= \langle \phi_\beta \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top + \langle \phi_\beta \boldsymbol{\eta}^\top \rangle \\
&= \langle \phi_\beta \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top
\end{aligned} \tag{3.15}$$

where, as in the explicit layered case, no statistical correlation exists between the pupil-plane phase vector and the measurement noise, and  $\langle \phi_\beta \boldsymbol{\eta}^\top \rangle = 0$ . Using Eqs. 3.14 – (3.15) the spatio-angular MMSE reconstructor becomes

$$\hat{\phi}_\beta \triangleq \langle \phi_\beta \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top (\mathbf{\Gamma} \langle \phi_\alpha \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top + \langle \boldsymbol{\eta} \boldsymbol{\eta}^\top \rangle)^{-1} \mathbf{s}_\alpha. \tag{3.16}$$

The simplified notation for the direction-specific phase covariance in the pupil-plane,  $\Sigma_{\beta, \alpha} = \langle \phi_\beta \phi_\alpha^\top \rangle$  will be used in later sections.

### 3.2.3 Equivalence of static MMSE reconstructors

Starting from the equations for the explicit reconstructor given in Eqs. 3.8-3.13, it can be shown that it is mathematically equivalent to the pupil-plane MMSE reconstructor, given that the solution is found for the same minimum pupil-integrated residual phase variance cost-functional from Eq. (3.1).

*Proof.* Proceed by identification. Let

$$\mathbf{P}_\beta \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle \mathbf{P}_\alpha^\top = \langle \phi_\beta \phi_\alpha^\top \rangle \tag{3.17}$$

be the pupil-plane phase angular covariance matrix between the m- $\boldsymbol{\beta}$  and n- $\boldsymbol{\alpha}$  directions and

$$\mathbf{P}_\alpha \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle \mathbf{P}_\alpha^\top = \langle \phi_\alpha \phi_\alpha^\top \rangle \tag{3.18}$$

be the covariance of phase in all directions ' $\boldsymbol{\alpha}'$ '. The latter is an  $n \times n$  matrix whereas  $\langle \boldsymbol{\varphi} \boldsymbol{\varphi}^\top \rangle$  is an  $n_L \times n_L$  block-diagonal matrix since layers in the atmosphere are considered independent. Hence, the estimated phase in the pupil becomes

$$\hat{\phi}_\beta = \langle \phi_\beta \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top (\mathbf{\Gamma} \langle \phi_\alpha \phi_\alpha^\top \rangle \mathbf{\Gamma}^\top + \langle \boldsymbol{\eta} \boldsymbol{\eta}^\top \rangle)^{-1} \mathbf{s}_\alpha \tag{3.19}$$

which is the solution one would otherwise obtain by directly applying the MMSE linear estimator. Thus, the *Spatio-Angular reconstructor* is a particular case of the *Explicit-tomography MMSE reconstructor*, applicable when one is only interested in the cumulative phase estimate in the pupil-plane.  $\square$

### 3.2.4 Overcoming temporal lag error

As discussed above, sky coverage can be improved by either making better use of every single photon (more efficient wave-front sensing), or by relaxing the temporal lag error constraint which allows for longer integration times if the error can be partly overcome by temporal prediction of the disturbances.

Lag error is intrinsic to any AO system due to the discrete nature of the measurements, their processing and correction. In Raven, a fixed system delay of 3ms has been allocated for camera read-out, wavefront reconstruction and DM actuation (the minimum delay in readout from the cameras is 2ms). The maximum frame rate of Raven will be 500 Hz (integration time of 2ms) but can and will be reduced by up to a factor of 10 according to the magnitude of the GSs being used. There are therefore two delay scenarios: The integration time is less than the system delay, in which case there are two frames of delay, or the integration time is greater than (or equal to) the system delay, in which case there is one frame plus a fraction of a frame of delay. Both scenarios are outlined in Fig. 3.1.

If we consider that the current measurement,  $\mathbf{s}_k$ ,

$$\mathbf{s}_k = \mathbf{\Gamma}\bar{\phi}_k + \boldsymbol{\eta}_k \quad (3.20)$$

is the result of a finite detector integration period over the previous sampling period, we can consider the measurement to be an average over that integration period and make the following definition:

$$\bar{\phi}_k = \frac{1}{T_s} \int_{(k-1)T_s}^{kT_s} \phi(t) dt. \quad (3.21)$$

This way, the measurement  $\mathbf{s}_k$  corresponds to the time at the termination of the exposure where the phase  $\phi_k$  was collected. The averaging is implied from now on.

All combined, the WFS data require several milliseconds to read out and compute which stresses the importance of a predictive capability. The lag error definition

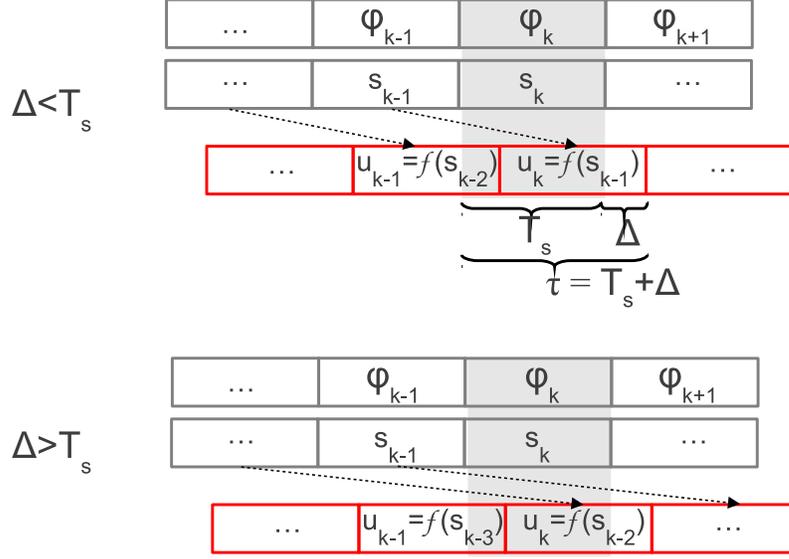


Figure 3.1: Temporal diagrams. Top:  $\text{Int}(\Delta/T_s) = 0$ , Commands  $u_k$  are conditioned to measurement  $s_{k-1}$ . Bottom:  $T_s < \Delta < 2T_s$ ;  $u_k$  is conditioned to  $s_{k-2}$ .

adopted here is the sum of the integration-time plus the pure-delay error accounting for the real-time processing of measurements and computation of DM commands.

An intermediate step between the static MMSE reconstructors and the spatially and temporally optimal Kalman filtering-based recursive LQG controller is temporal prediction. Temporal prediction is a step that must be developed anyway in order to implement the LQG controller. Therefore, an examination of its potential outside of this algorithm was undertaken with the goal of achieving improved performance without adding a large amount of complexity to the real-time data processing pipeline. Rewriting the spatial minimization criterion given in Eq. 3.1 to include time, we would like to minimize

$$\mathbf{R} = \arg \min_{\mathbf{R}'} \langle \|\phi(\boldsymbol{\rho}, t + T_s) - \mathbf{R}\mathbf{s}_\alpha(t)\|_{L_2(\Omega)}^2 \rangle. \quad (3.22)$$

Using Taylor's frozen flow hypothesis, rewrite  $\phi(\boldsymbol{\rho}, t + T_s) = \phi(\boldsymbol{\rho} - T_s\mathbf{v}, t)$  and assume there is a linear operator which translates the phase,  $\phi(\boldsymbol{\rho} - T_s\mathbf{v}, t) = \mathcal{A}\phi(\boldsymbol{\rho}, t)$ . The primary assumptions are that we have good knowledge of the wind velocity,  $\mathbf{v}$ , in each layer of the atmosphere, that each layer of the atmosphere is independent from the others, and that the atmosphere can be described by a Kolmogorov or von Kármán spectrum [34], [63]. The minimization can once again be expressed in an

exclusively spatial domain,

$$\mathbf{R} = \arg \min_{\mathbf{R}'} \langle \|\mathcal{A}\phi(\boldsymbol{\rho}) - \mathbf{R}\mathbf{s}_\alpha\|_{L_2(\Omega)}^2 \rangle. \quad (3.23)$$

Assuming the phase is estimated in a specific direction vector,  $\boldsymbol{\beta}$ , as before, it can be seen that the predicted phase estimate can be obtained by applying the translation directly to the static reconstructor,

$$\phi_\beta(t + T_s) = \mathcal{A}\Sigma_{\phi_\beta, \mathbf{s}_\alpha} \Sigma_{\mathbf{s}_\alpha}^{-1} \mathbf{s}_\alpha(t). \quad (3.24)$$

A full treatment of this problem can be found in [64] Prediction models are developed for both the modal and zonal bases in Secs 3.3.2 and 3.4.2 respectively.

### 3.2.5 Linear Quadratic Gaussian Controller

The phase estimation techniques described above are sub-optimal reconstructors. The well established optimal solution, which minimizes the function given in Eq. 3.23, corresponds to a discrete-time LQG controller [65]. The LQG controller is appropriate for use in AO systems as it can take advantage of the stochastic nature of the atmosphere. It is particularly well-suited for tomographic AO as both spatial and temporal errors make significant contributions to the overall performance. Primary applications of the LQG controller in tomographic AO systems have focussed on MCAO (CL systems) [66, 67, 68], however the formalism is the same for both closed and open loop and no adjustments are required in order to apply it to an MOAO system. The format of the controller is also very flexible, allowing for the incorporation of asynchronous control as well as vibration cancellation. The purpose of an LQG controller is to generate an optimal stochastic controller with respect to, in the case of AO, residual phase error. The system is assumed to be modelled by the linear stochastic difference equation and state feedback equation,

$$\mathbf{x}_{k+1} = \mathcal{A}_d \mathbf{x}_k + \mathcal{B}_d \mathbf{u}_k + \mathbf{V}_k \quad (3.25)$$

$$\mathbf{s}_k = \mathcal{C}_d \mathbf{x}_k + \mathcal{D}_d \mathbf{u}_k + \boldsymbol{\eta}_k, \quad (3.26)$$

where  $\mathbf{x}_k$  is the input state. The matrix  $\mathcal{A}_d$  is the state transition matrix; it relates the state at time  $k$  to the state at time  $k + 1$ . The matrix  $\mathcal{C}_d$  is the measurement model which relates the state to the measurement,  $\mathbf{s}_k$ . The measurements are slopes

in the case of a SH-WFS AO system and  $\mathcal{C}_d$  is a linear operator, generally containing the measurement model  $\mathbf{\Gamma}$  used in previous sections to take a selected portion of the state from phase space to slope space. The quantities  $\mathbf{V}, \boldsymbol{\eta}$  are independent zero-mean Gaussian white noise signals representing process noise and measurement noise respectively; the covariances of these two noise processes are represented by  $\mathbf{Q}$  and  $\mathbf{R}$ . In the case of an MOAO system, which is open loop, neither the state nor the current measurement have any dependence on the the feedback term,  $\mathbf{u}_k$ , therefore the terms  $\mathcal{B}_d, \mathcal{D}_d$  are zero.

When discussing the state estimate,  $\hat{\mathbf{x}}_k$ , it is important to make the following distinction:  $\hat{\mathbf{x}}_{k|k-1}$  refers to the *a priori* state estimate which is the state estimate at time  $k$  given knowledge of the process prior to step  $k$ .  $\hat{\mathbf{x}}_{k|k}$  is the *a posteriori* state estimate which is the state estimate at step  $k$  given measurement  $\mathbf{s}_k$ . We can now define terms for the state estimate error in each case:

$$\mathbf{e}_{k|k-1} \equiv \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \quad (3.27)$$

$$\mathbf{e}_{k|k} \equiv \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}, \quad (3.28)$$

and the covariance of the *a posteriori* error estimate can be expressed,

$$\boldsymbol{\Sigma}_k = \langle \mathbf{e}_{k|k} \mathbf{e}_{k|k}^T \rangle. \quad (3.29)$$

The goal is to find an equation that computes an *a posteriori* state estimate,  $\hat{\mathbf{x}}_{k|k}$ , as a linear combination of an *a priori* estimate,  $\hat{\mathbf{x}}_{k|k-1}$  and the weighted difference between an actual measurement,  $\mathbf{s}_k$ , and the measurement prediction, made at the previous time step,

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathcal{M}_k(\mathbf{s}_k - \hat{\mathbf{s}}_{k|k-1}), \quad (3.30)$$

which minimizes the *a posteriori* error. The measurement estimate is expressed,

$$\hat{\mathbf{s}}_{k|k-1} = \mathcal{C}_d \hat{\mathbf{x}}_{k|k-1}, \quad (3.31)$$

and the weighting is carried out by the matrix  $\mathcal{M}_k$ . The term  $\mathbf{s}_k - \hat{\mathbf{s}}_{k|k-1}$  is referred to as the measurement innovation and reflects the discrepancy between the predicted measurement and actual measurement at time  $k$ . The matrix  $\mathcal{M}_k$  is the Kalman gain

which minimizes the *a posteriori* error covariance  $\Sigma_k$ ,

$$\mathcal{M}_k = \Sigma_k \mathcal{C}_d^\top (\mathcal{C}_d \Sigma_k \mathcal{C}_d^\top + \mathbf{R})^{-1}. \quad (3.32)$$

Substituting Eq. 3.32 into Eq. 3.30 and subsequently into Eq. 3.28, the expectation in Eq. 3.29 can be developed. The result of this operation is the discrete, time-invariant Riccati difference equation [65],

$$\Sigma_{k+1} = \mathcal{A}_d \Sigma_k \mathcal{A}_d^\top + \mathbf{Q} - \mathcal{A}_d \Sigma_k \mathcal{C}_d^\top (\mathcal{C}_d \Sigma_k \mathcal{C}_d^\top + \mathbf{R})^{-1} \mathcal{C}_d \Sigma_k \mathcal{A}_d^\top, \quad (3.33)$$

which is solved recursively for  $\Sigma_k$ , called  $\Sigma_\infty$  as it approaches convergence. In the case of AO, the WF error must be minimized for long exposure images, thus this asymptotic solution can be used without loss of performance. The gain,  $\mathcal{M}_\infty$  is expressed as,

$$\mathcal{M}_\infty = \Sigma_\infty \mathcal{C}_d^\top (\mathcal{C}_d \Sigma_\infty \mathcal{C}_d^\top + \mathbf{R})^{-1}. \quad (3.34)$$

In practice, there are three real-time steps to the recursive filter, assuming that the amount of delay in the system is a full sample period:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathcal{M}_\infty (\mathbf{s}_k - \mathcal{C}_d \hat{\mathbf{x}}_{k|k-1}) \\ \hat{\mathbf{x}}_{k+1|k} &= \mathcal{A}_d \hat{\mathbf{x}}_{k|k}. \\ \mathbf{u}_k &= (\mathbf{IM})^\dagger \mathbf{T} \hat{\mathbf{x}}_{k+1|k} \end{aligned} \quad (3.35)$$

The first is the computation of the measurement estimate from the predicted phase of the previous step combined with the estimation of the state, as in Eq. 3.30. The second is the prediction step, or state transition, which is the same as the pure prediction described in Sec. 3.2.4. In this notation, the matrix  $\mathcal{A}_d$  is a composite matrix which transitions the state using a linear combination of the state elements according to the particular prediction model. The LQG controller is developed for several different prediction models in modal and zonal bases in Sections 3.3.4 and 3.4.4. The final step is the computation of DM commands from the estimated, predicted state via conversion from phase space to command space. For practical reasons, the phase is mapped back onto slopes in the  $\beta$  directions; this is due to the DM calibration available from the system. In Raven, the science DMs (SDMs) are calibrated via an interaction matrix between them and the CL-WFS on each science channel. The easiest DM fitting step therefore goes from slopes to commands, so it is convenient

to estimate  $\hat{\mathbf{s}}_\beta$

### Non-integer frame delay

The assumption depicted by Eqs. 3.35 is that the amount of delay in the system is a full sample period. As stated above, in Raven (or in any instrument), this may not be the case. With the LQG controller, timing is of key importance; care must be taken to apply the predicted commands at the time step for which they were conditioned. The particular application to MOAO systems, specifically the open-loop tomography, allows for the construction of a LQG model that easily admits asynchronous timing cases. To do so one assumes the measurements are available at the end of the integration step, regardless of the actual delay. One affects thus the whole delay to the commands instead of the measurements.

In Eq. 3.35, the measurement estimate is  $\hat{\mathbf{s}}_{k|k-1} = \mathbf{\Gamma}\hat{\boldsymbol{\phi}}_{k|k-1}$ , and the state vector,  $\mathbf{x}$  contains  $N_k$  phase estimates,

$$\hat{\mathbf{x}}_{k|k} = \left[ \hat{\boldsymbol{\phi}}_{k|k}, \hat{\boldsymbol{\phi}}_{k-1|k}, \dots, \hat{\boldsymbol{\phi}}_{k-N_k|k} \right]^\top, \quad (3.36)$$

Applying the prediction step to get  $\hat{\mathbf{x}}_{k+1|k}$  and using the assumption of a fractional number of frame delays for readout and processing, we would like to compute the DM commands from the current measurement  $\mathbf{s}_k$  at the moment it will actually be applied in the system, *i.e.* at time  $k + \Delta$ , for fractional delays either greater or smaller than the integration time.

The controller is applied in real-time by computing, at iteration  $k$  for either case  $\Delta \geq 1$  or  $\Delta \leq 1$ ,  $\Delta \in \mathbb{Q}^+$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathcal{M}_\infty (\mathbf{s}_k - \mathcal{C}_d \hat{\mathbf{x}}_{k|k-1}) \quad (3.37a)$$

$$\hat{\mathbf{x}}_{k+1|k} = \mathcal{A}_d \hat{\mathbf{x}}_{k|k} \quad (3.37b)$$

$$\begin{aligned} \hat{\mathbf{x}}_{k+1+\Delta|k} &= \mathcal{A}_\Delta \hat{\mathbf{x}}_{k+1|k} \\ \mathbf{u}_k &= (\mathbf{I}\mathbf{M})^\dagger \mathbf{\Gamma} \hat{\mathbf{x}}_{k+1+\Delta|k} \end{aligned} \quad (3.37c)$$

where  $\mathcal{A}_\Delta$  is the matrix that overcomes the pure lag delay and is not included in the recursion.

It will be shown in Sec. 3.3.3 that the greater the system lag, the more benefit can be derived from carrying out temporal prediction (see Fig. 3.7). With this in mind, and looking forward to simulation results and laboratory measurements to be

shown in Chaps. 4 and 5 respectively, it is likely that the system will be operated in the scenario of  $\Delta < T_s$  when prediction or LQG control is being carried out.

### 3.3 Modal reconstruction

When a modal basis is used to develop the reconstructor, the general measurement model given in Eq. 2.52 can be specified as,

$$\mathbf{s} = \mathbf{\Gamma} \mathbf{z}_c + \boldsymbol{\eta}, \quad (3.38)$$

where  $\mathbf{\Gamma}$  is the linear operator that transforms Zernike coefficients to slopes and  $\mathbf{z}_c$  is a vector of Zernike coefficients representing the phase.

#### 3.3.1 Static Reconstructors in Zernike Space

Recalling the MMSE solution from Eq. 3.16,

$$\hat{\boldsymbol{\phi}}_\beta \triangleq \langle \boldsymbol{\phi}_\beta \boldsymbol{\phi}_\alpha^\top \rangle \mathbf{\Gamma}^\top (\mathbf{\Gamma} \langle \boldsymbol{\phi}_\alpha \boldsymbol{\phi}_\alpha^\top \rangle \mathbf{\Gamma}^\top + \langle \boldsymbol{\eta} \boldsymbol{\eta}^\top \rangle)^{-1} \mathbf{s}_\alpha \quad (3.39)$$

It is clear that an expression for the covariance of the atmospheric phase in the Zernike basis is required. The angular pupil-plane covariance matrix between any Zernike polynomial is analytically computed from [69] for the infinite outer-scale case of turbulence

$$\langle a_i(0), a_i(\xi) \rangle = 3.895 \left( \frac{D}{r_0} \right)^{\frac{5}{3}} \frac{\int_0^{h_{\max}} C_n^2(h) I_{ij} \left( \frac{\xi h}{R} \right) dh}{\int_0^\infty C_n^2(h) dh} \quad (3.40)$$

with  $D = 2R$  the telescope diameter,  $r_0$  the Fried parameter,  $h$  the altitude above the telescope,  $\xi$  the angle between the pupils over which the Zernike polynomials are defined,  $C_n^2$  the atmospheric vertical profile and,

$$\begin{aligned} I_{ij}(x) = & (-1)^{\frac{n_1+n_2-m_1-m_2}{2}} \sqrt{(n_1+2)(n_2+1)} \times \\ & \left[ K_{1,2}^+ \int_0^\infty \kappa^{-\frac{14}{3}} J_{n_1+1}(2\pi\kappa) J_{n_2+1}(2\pi\kappa) J_{m_1+m_2}(2\pi\kappa x) d\kappa \right. \\ & \left. + K_{1,2}^- \int_0^\infty \kappa^{-\frac{14}{3}} J_{n_1+1}(2\pi\kappa) J_{n_2+1}(2\pi\kappa) J_{|m_1-m_2|}(2\pi\kappa x) d\kappa \right] \end{aligned} \quad (3.41)$$

with  $K_{1,2}^+$  and  $K_{1,2}^-$  coefficients that depend on  $m_i$  and  $n_i$ . For the general case, take  $\xi$  to be the angular difference of  $\alpha_i$  and  $\alpha_j$ .

The cross-correlation functions have been extended for the finite outer-scale case in [70] and later extensively used and generalized in [71, 62]. The layered spatial covariance matrix is a block-diagonal matrix (layers are independent) and can be found in [6] for the infinite outer scale case and in [72] for the finite case.

The measurement model,  $\mathbf{\Gamma}$ , is a concatenation of derivatives of Zernike polynomials in  $x$  and  $y$  sampled over a grid with resolution equal to the number of WFS lenslets. The derivatives of Zernike polynomials can be represented as a linear combination of Zernike polynomials, as given by [6],

$$\nabla Z_j = \sum_{j'} \gamma_{jj'} Z_{j'}. \quad (3.42)$$

The coefficients,  $\gamma_{jj'}$  can be expressed in rectangular coordinates, resulting in an  $x$  and  $y$  derivative,

$$\gamma_{jj'}^x = \int d^2 \rho Z_{j'} \frac{dZ_j}{dx} \quad (3.43)$$

$$\gamma_{jj'}^y = \int d^2 \rho Z_{j'} \frac{dZ_j}{dy} \quad (3.44)$$

It was shown in Sec. 3.2.3 that explicit layered and spatio-angular formulations of the static MMSE reconstructor are equivalent. In a modal basis such as Zernike space, this equivalence is only strictly true when the number of Zernike modes considered when computing the propagation matrices,  $\mathbf{P}_\alpha, \mathbf{P}_\beta$ , and the phase covariance,  $\langle \varphi \varphi^T \rangle$  of the explicit-tomography reconstructor goes to infinity. The number of reconstructed modes is truncated according to the order of the system (through out this work,  $n = 9$  radial orders are used as all WFSs are 10x10), however the projection of all modes is technically required to eliminate aliasing effects. In practice, a sufficiently large number of modes can be determined which reduces the error in the first 9 radial orders to a negligible amount. As has been shown, the expressions,  $\mathbf{P}_\alpha \langle \varphi \varphi^T \rangle \mathbf{P}_\alpha^T$  and  $\langle \phi_\alpha \phi_\alpha^T \rangle$  should give identical results. By performing the matrix multiplications for a specific case with a single atmospheric layer and comparing the diagonals of the two resulting matrices, Fig. 3.2 shows that as the number of modes used to compute the left-hand side of Eq. 3.18 is increased, its diagonal approaches that of the right-hand side, which is computed analytically.

The propagation matrices,  $\mathbf{P}_\beta, \mathbf{P}_\alpha$ , are the modal projection of Zernike polynomi-

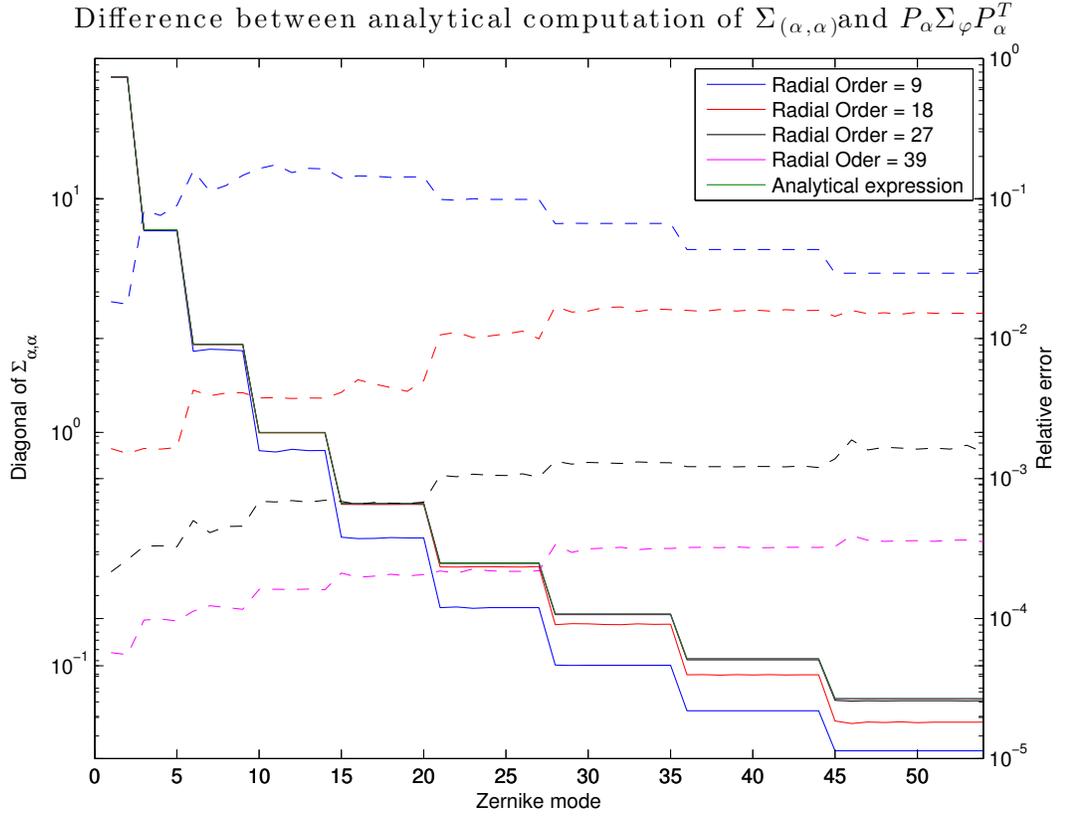


Figure 3.2: Aliasing error in the computation of explicit layered covariance matrix,  $\mathbf{P}_\alpha \Sigma_\varphi \mathbf{P}_\alpha^T$ , is reduced with increasing radial order.

als onto a resized pupil which is displaced in  $x$  and  $y$ . They are computed analytically using Noll indexing based on the function *TransformC* developed by Lundstrom and Unsbo and available in [73].

### 3.3.2 Predictive MMSE in Zernike space

Several temporal prediction models are developed for the modal basis reconstructor and a theoretical analysis of their ability to reduce temporal lag error is presented here. Based on this analysis, only two were selected for comparative study in simulation.

#### Auto-Regressive models of first, second and third order

Auto-Regressive (AR) models are relatively coarse models, typically used for prediction when embedded in the reconstructor and implemented in a dynamic controller such as Kalman filtering. They can be used for the off-line computation of low order

optimal gains. The Zernike-space AR models assume that each Zernike mode behaves independently, thus generating diagonal prediction matrices.

An AR model of order  $n$  is defined by the recursion,

$$\boldsymbol{\varphi}_{k+1} = f(\boldsymbol{\varphi}_k, \dots, \boldsymbol{\varphi}_{k-n-1}) + \boldsymbol{\varepsilon}_k, \quad (3.45)$$

where  $f(\dots)$  is a linear function and  $\boldsymbol{\varepsilon}_k$  is called the excitation noise; it is a Gaussian-distributed spectrally white zero-mean random sequence with variance such that the output variance is conserved and is constrained to be equal to the value prescribed by the Kolmogorov or von Kármán turbulence models. The covariance of  $\boldsymbol{\varepsilon}_k$  can be derived for each AR model from the expression of the specified model.

A first order AR model (AR1) employs the current estimate only,

$$\boldsymbol{\varphi}_{k+1}^{\text{AR1}} = \mathbf{A}_{\text{AR1}} \boldsymbol{\varphi}_k^{\text{AR1}} + \boldsymbol{\varepsilon}_k^{\text{AR1}} \quad (3.46)$$

Computing the covariance of both sides leads directly to an expression for the excitation noise covariance,

$$\begin{aligned} \langle \boldsymbol{\varphi}_{k+1}^{\text{AR1}} (\boldsymbol{\varphi}_{k+1}^{\text{AR1}})^{\text{T}} \rangle &= \langle (\mathbf{A}_{\text{AR1}} \boldsymbol{\varphi}_k^{\text{AR1}} + \boldsymbol{\varepsilon}_k^{\text{AR1}}) (\mathbf{A}_{\text{AR1}} \boldsymbol{\varphi}_k^{\text{AR1}} + \boldsymbol{\varepsilon}_k^{\text{AR1}})^{\text{T}} \rangle \\ \boldsymbol{\Sigma}_{\boldsymbol{\varphi}}^{\text{AR1}} &= \mathbf{A}_{\text{AR1}} \boldsymbol{\Sigma}_{\boldsymbol{\varphi}}^{\text{AR1}} \mathbf{A}_{\text{AR1}}^{\text{T}} + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}_k}^{\text{AR1}} \\ \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}_k}^{\text{AR1}} &= \boldsymbol{\Sigma}_{\boldsymbol{\varphi}}^{\text{AR1}} - \mathbf{A}_{\text{AR1}} \boldsymbol{\Sigma}_{\boldsymbol{\varphi}}^{\text{AR1}} \mathbf{A}_{\text{AR1}}^{\text{T}}. \end{aligned} \quad (3.47)$$

Note that

$$\langle \boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^{\text{T}} \rangle = \langle \boldsymbol{\varphi}_{k+1} \boldsymbol{\varphi}_{k+1}^{\text{T}} \rangle = \boldsymbol{\Sigma}_{\boldsymbol{\varphi}} = \boldsymbol{\Sigma}_{\boldsymbol{\varphi}}^{\text{AR1}}, \quad (3.48)$$

which, due to stationarity, loses its temporal dependence. In the AR1 case, the prediction model,  $\mathbf{A}_{\text{AR1}}$ , is a diagonal matrix of independent modal coefficients. The method for estimating these coefficients is discussed below.

Using the second order AR model (AR2), the phase in each atmospheric layer can be estimated as a linear combination of the phase at the two previous time-steps plus process noise,

$$\boldsymbol{\varphi}_{k+1}^{\text{AR2}} = \mathbf{A}_{\text{AR2}} \boldsymbol{\varphi}_k^{\text{AR2}} + \mathbf{B}_{\text{AR2}} \boldsymbol{\varphi}_{k-1}^{\text{AR2}} + \boldsymbol{\varepsilon}_k^{\text{AR2}} \quad (3.49)$$

Assuming the coefficients of  $\mathbf{A}_{\text{AR2}}$  and  $\mathbf{B}_{\text{AR2}}$  are known, and imposing  $\boldsymbol{\Sigma}_{\boldsymbol{\varphi}}^{\text{AR2}} = \boldsymbol{\Sigma}_{\boldsymbol{\varphi}}$  as before, compute the expectation as in the AR1 case and solve for the process noise

covariance,  $\Sigma_\epsilon^{AR2}$ ,

$$\Sigma_\epsilon^{AR2} = \Sigma_\varphi - \mathbf{A}_{AR2} \Sigma_\varphi \mathbf{A}_{AR2}^T - \mathbf{B}_{AR2} \Sigma_\varphi \mathbf{B}_{AR2}^T - \mathbf{A}_{AR2} \Sigma_{1\tau}^{AR2} \mathbf{B}_{AR2}^T - \mathbf{B}_{AR2} \Sigma_{1\tau}^{AR2} \mathbf{A}_{AR2}^T \quad (3.50a)$$

$$\Sigma_{1\tau}^{AR2} = \langle \varphi_{k+1}^{AR2} (\varphi_k^{AR2})^T \rangle \triangleq \langle \varphi_k^{AR2} (\varphi_{k-1}^{AR2})^T \rangle = (\mathbf{I} - \mathbf{B}_{AR2})^{-1} \mathbf{A}_{AR2} \Sigma_\varphi \quad (3.50b)$$

Noting that, again, due to stationarity,

$$\langle \varphi_{k+1}^{AR2} (\varphi_k^{AR2})^T \rangle = \langle \varphi_k^{AR2} (\varphi_{k+1}^{AR2})^T \rangle = \Sigma_{1\tau}^{AR2}. \quad (3.51)$$

$\Sigma_{1\tau}^{AR2}$  is a one-step modal spatio-temporal cross-covariance matrix between Zernike modes at any times  $k+1$  and  $k$  separated by a single time step. Use Eq. 3.49,  $\varphi_k^T$ , and the expression for  $\Sigma_{1\tau}^{AR2}$  in Eq. 3.51, and take the expectation,

$$\langle \varphi_{k+1}^{AR2} (\varphi_k^{AR2})^T \rangle = \langle (\mathbf{A}_{AR2} \varphi_k^{AR2} + \mathbf{B}_{AR2} \varphi_{k-1}^{AR2} + \epsilon_k^{AR2}) (\varphi_k^{AR2})^T \rangle, \quad (3.52)$$

$$\langle \varphi_{k+1}^{AR2} (\varphi_k^{AR2})^T \rangle = \mathbf{A}_{AR2} \langle \varphi_k^{AR2} (\varphi_k^{AR2})^T \rangle + \mathbf{B}_{AR2} \langle \varphi_{k-1}^{AR2} (\varphi_k^{AR2})^T \rangle, \quad (3.53)$$

$$\Sigma_{1\tau}^{AR2} = \mathbf{A}_{AR2} \Sigma_\varphi + \mathbf{B}_{AR2} \Sigma_{1\tau}^{AR2}. \quad (3.54)$$

Solving for  $\Sigma_{1\tau}^{AR2}$  gives,

$$\Sigma_{1\tau}^{AR2} = \frac{\mathbf{A}_{AR2}}{1 - \mathbf{B}_{AR2}} \Sigma_\varphi. \quad (3.55)$$

This matrix can also be computed by noting that the temporal covariance between adjacent time steps is equivalent to the spatial covariance between points  $x_0$  and  $x_1 = vT_s$  for a given wind-speed and sampling time. This is exactly the computation described by Eqs. 3.40 and 3.41 where  $a_i(0), a_i(\xi)$  is the angular separation between  $x_0$  and  $x_1$ .

For completeness, the third order model (AR3) is developed,

$$\varphi_{k+1}^{AR3} = \mathbf{A}_{AR3} \varphi_k^{AR3} + \mathbf{B}_{AR3} \varphi_{k-1}^{AR3} + \mathbf{C}_{AR3} \varphi_{k-2}^{AR3} + \epsilon_k^{AR3} \quad (3.56)$$

the noise covariance matrix is found from

$$\begin{aligned} \Sigma_\epsilon^{AR3} = & \Sigma_\varphi - \mathbf{A}_{AR3} \Sigma_\varphi \mathbf{A}_{AR3}^T - \mathbf{B}_{AR3} \Sigma_\varphi \mathbf{B}_{AR3}^T - \mathbf{C}_{AR3} \Sigma_\varphi \mathbf{C}_{AR3}^T \\ & - \mathbf{A}_{AR3} \Sigma_{1\tau}^{AR3} \mathbf{B}_{AR3}^T - \mathbf{B}_{AR3} \Sigma_{1\tau}^{AR3} \mathbf{A}_{AR3}^T - \mathbf{B}_{AR3} \Sigma_{1\tau}^{AR3} \mathbf{C}_{AR3}^T - \mathbf{C}_{AR3} \Sigma_{1\tau}^{AR3} \mathbf{B}_{AR3}^T \\ & - \mathbf{A}_{AR3} \Sigma_{2\tau}^{AR3} \mathbf{C}_{AR3}^T - \mathbf{C}_{AR3} \Sigma_{2\tau}^{AR3} \mathbf{A}_{AR3}^T \end{aligned} \quad (3.57a)$$

where  $\Sigma_{1\tau}^{\text{AR3}}$  is again the 1-step modal cross-covariance matrix this time for the AR3 model and  $\Sigma_{2\tau}^{\text{AR3}} \triangleq \langle \varphi_{k+1}^{\text{AR3}}(\varphi_{k-1}^{\text{AR3}})^{\text{T}} \rangle \triangleq \langle \varphi_k^{\text{AR3}}(\varphi_{k-2}^{\text{AR3}})^{\text{T}} \rangle$  is a two-step modal spatio-temporal cross-covariance matrix, with

$$\Sigma_{1\tau}^{\text{AR3}} = (\mathbf{B}_{\text{AR3}} + \mathbf{C}_{\text{AR3}}\mathbf{A}_{\text{AR3}} + \mathbf{C}_{\text{AR3}}^2)^{-1} (\mathbf{A}_{\text{AR3}} + \mathbf{C}_{\text{AR3}}\mathbf{B}_{\text{AR3}}) \Sigma_{\varphi} \quad (3.57b)$$

$$\Sigma_{2\tau}^{\text{AR3}} = \mathbf{B}_{\text{AR3}}\Sigma_{\varphi} + (\mathbf{A}_{\text{AR3}} + \mathbf{C}_{\text{AR3}}\mathbf{B}_{\text{AR3}}) \Sigma_{1\tau}^{\text{AR3}} \quad (3.57c)$$

### AR model parameter identification

To carry out temporal prediction, we must identify the coefficients of the prediction matrices  $\mathcal{A}$ ,  $\mathcal{B}$ , and up to  $\mathcal{N}$  (for arbitrarily high ordered AR models). This must be done such that the model has the same temporal auto-correlation as the data (within a reasonable temporal horizon). A practical method to identify the coefficients is to fit the initial  $T_{\text{fit}}$  seconds of the temporal auto-correlation function of each and every mode as previously suggested in [74]. An alternative is to match the decorrelation at the coherence time of each mode [66]. For AR1 models, both strategies lead to roughly the same modal decorrelation functions and can therefore be used interchangeably. For higher order models (AR2 and AR3), fitting the initial  $T_{\text{fit}}$  seconds leads to the overall best prediction performances, although the existence of a general  $T_{\text{fit}}$  is debatable and subject to optimization.

In general, we would like to determine some parameters  $\mathbf{a}, \mathbf{b}, \dots, \mathbf{n}$  of a model of the modal atmospheric temporal covariance function,  $\mathcal{C}(t)_{\text{model}}$ , such that the error between it and the analytical modal atmospheric temporal covariance function,  $\mathcal{C}(t)$  is minimized mode-by-mode for each Zernike in each layer over a given temporal horizon,

$$(\mathbf{a}, \dots, \mathbf{n}) = \arg \min_{\mathbf{a}, \dots, \mathbf{n}} \int_0^{T_{\text{fit}}} |\mathcal{C}(t) - \mathcal{C}(t)_{\text{model}}|^2 dt. \quad (3.58)$$

The current estimated state of the atmosphere (estimated layer altitudes, wind-speeds, and power) can be used to compute an expression for the current temporal covariance matrix of Zernike modes in the atmosphere. According to the Wiener-Khinchine theorem [75, 36], the theoretical temporal autocorrelation of Zernike modes can be computed from the normalized Fourier transform of the temporal PSD,

$$\mathcal{C}(t) = \mathcal{F}\{\mathbf{W}_t(\nu)\}. \quad (3.59)$$

In this case  $\mathcal{C}(t)$  is the normalized Fourier transform of the temporal PSD,  $\mathbf{W}_t^j$ , of

each mode in each layer. For a given layer with wind velocity  $\mathbf{v}$ , the temporal PSDs of the modes can be computed analytically, as outlined in [76], as the integral of the product of the spatial power spectrum,  $\mathbf{W}_\phi(\boldsymbol{\omega})$  and the squared modulus of a spatial function,  $M_G(\mathbf{r})$  defined on the pupil and related to the temporal domain via the Taylor frozen flow hypothesis [77]. The spatial PSD for a given layer with fractional  $r_0, f_{r_0}$  is given in Eq. 2.35. We can model the temporal PSD as the square modulus of the continuous time transfer function of the ARn model given in Eq. 3.45. Thus the modelled temporal autocorrelation of each mode is computed from the Fourier transform of the square modulus of that transfer function,

$$\mathcal{C}(t)_{model} = \mathcal{F}\{|H(s)|^2\}. \quad (3.60)$$

Therefore it is the coefficients of the continuous time transfer function we wish to identify in order to minimize Eq. 3.58. For the ARn system represented by the recursion in Eq. 3.45, the continuous time transfer function of mode  $j$  is,

$$H_j(s) = \frac{1}{s^n - a_j s^{n-1} - b_j s^{n-2} - \dots - n_j}. \quad (3.61)$$

The coefficients,  $[a_j, b_j, \dots, n_j]$  are the continuous time coefficients for mode  $j$  which are diagonal elements of the matrices  $[\mathcal{A}, \mathcal{B}, \dots, \mathcal{N}]$ . This expression can be factored and written in terms of the poles of the transfer function, a form that will be useful for converting to a discrete model for practical applications,

$$H_j(s) = \frac{1}{(s - p_1)(s - p_2)(\dots)(s - p_n)}. \quad (3.62)$$

Specific examples of the derivation of the continuous transfer function are given here for AR1 and AR2 models. Starting from Eq. 3.46 for an AR1 model, the discrete problem can be recast to the continuous domain by replacing the samples  $[k, k+1, \dots, k+n]$  by  $[t, t+\Delta, \dots, t+n\Delta]$ . As  $\Delta$  becomes small, the expression can be written in terms of its temporal partial derivatives,

$$\dot{\varphi}(t) - \mathcal{A}\varphi(t) = \epsilon(t). \quad (3.63)$$

The discrete excitation noise is converted to continuous excitation noise as described in [68]. Recall that the phase vector is a set of Zernike modes, therefore we can take the Laplace transform of Eq. 3.63 mode-by-mode. Rearranging gives an expression

of the continuous transfer function for mode  $j$ ,

$$\Phi_j(s)s - a_j\Phi_j(s) = E(s) \quad (3.64)$$

$$H_j(s) = \frac{E(s)}{\Phi_j(s)} = \frac{1}{s - a_j}. \quad (3.65)$$

Similarly, for a second order model, the continuous representation of Eq. 3.49 can be written as a temporal differential equation,

$$\ddot{\varphi}(t) - \mathcal{A}\dot{\varphi}(t) - \mathcal{B}\varphi(t) = \epsilon(t). \quad (3.66)$$

Taking the Laplace transform of Eq. 3.66 mode-by-mode and rearranging once again gives an expression of the continuous transfer function of mode  $j$ ,

$$\Phi_j(s)s^2 - a_j\Phi_j(s)s - b_j\Phi_j(s) = E(s), \quad (3.67)$$

$$\Phi_j(s)(s^2 - a_js - b_j) = E(s), \quad (3.68)$$

$$H_j(s) = \frac{1}{s^2 - a_js - b_j}. \quad (3.69)$$

With a computed temporal covariance,  $\mathcal{C}(t)$ , and a model of the temporal covariance,  $\mathcal{C}(t)_{model}$ , we can now determine the coefficients  $[a_j, b_j, \dots, n_j]$ , such that the minimization problem in Eq. 3.58 is satisfied for each Zernike mode in each layer over a selected temporal horizon. Because the highest order AR model considered is AR3, and the sample times fall in the range of 2 to 20 milliseconds, the temporal horizon does not need to be larger than 0.1 seconds.

There are many angles from which to approach this parameter identification problem. The first attempt was a brute force two dimensional grid search using the built in Matlab function *fminbnd* (a bounded one dimensional search). This method works well enough for a one dimensional parameter search, but it quickly becomes a liability for higher dimensions as it suffers from extremely long computing time when using fine enough sampling required in order to achieve good accuracy.

### Broyden-Fletcher-Goldfarb-Shanno method

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method for parameter identification is of the quasi-Newton algorithm class for multidimensional optimization. The basic steps of a quasi-Newton algorithm are as follows [78]:

- **Step 1:**

Input an initial parameter vector,  $x_0$ , and initialize the tolerance  $\epsilon$ .

Set the iteration count to zero,  $k = 0$ , and approximate the second partial derivative matrix of the function as the identity,  $S_0 = I_n$ .

Compute the vector of first partial derivatives,  $g_0$ , at  $x_0$ .

- **Step 2:**

Set  $d_k = -S_k g_k$ .

Find  $\alpha_k$ , the value of  $\alpha$  that minimizes  $f(x_k + \alpha d_k)$ , using a one dimensional line search.

Set  $\delta_k = \alpha_k d_k$  and  $x_{k+1} = x_k + \delta_k$ .

- **Step 3:**

If  $\|\delta_k\| < \epsilon$ , output  $x^* = x_{k+1}$  and  $f(x^*) = f(x_{k+1})$ , and stop.

- **Step 4:**

Compute  $g_{k+1}$  and set  $\gamma_k = g_{k+1} - g_k$ .

Compute  $S_{k+1}$  using the BFGS updating formula (see Eq. 3.70).

Set  $k = k + 1$  and repeat from Step 2.

The BFGS updating formula is given by,

$$S_{k+1} = S_k + \left(1 + \frac{\gamma_k^T S_k \gamma_k}{\gamma_k^T \delta_k}\right) \frac{\delta_k \delta_k^T}{\gamma_k^T \delta_k} - \frac{\delta_k \gamma_k^T S_k + S_k \gamma_k \delta_k^T}{\gamma_k^T \delta_k}. \quad (3.70)$$

As  $k$  increases,  $S_k$  approaches the matrix of second partial derivatives of  $f(x_k)$  without requiring a closed form expression. In this case,  $f(x_k) = \sum_{t=0}^{T_{fit}} |\Sigma(t) - \Sigma(t)_{model}|^2$  (the continuous integral from Eq. 3.58 is replaced by a finely sampled time vector). In place of computing a closed form expression for the partial derivatives  $g_k$  of this function, a numerical approximation was used,

$$\frac{\partial f(x_k)}{\partial x_i} = \frac{f(x_1, \dots, x_i + \delta, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{\delta} \quad (3.71)$$

where  $\delta$  is a fixed small value, typically  $10^{-9} < \delta < 10^{-6}$ , and is not related to  $\delta_k$  in Eq. 3.70.

In the AR2 case, the results of the brute force parameter identification process was compared to the BFGS method described here. The two dimensional *fminbnd* search required several hours, placing it outside of the realm of practical implementation in any system, whereas the the BFGS method took less than 60 seconds. The resulting (discretized) coefficients of the first state matrix,  $\mathbf{A}_{AR2}$ , identified by each method for the first 45 Zernike modes are shown in Fig. 3.3 and shows that there is good agreement between the methods.

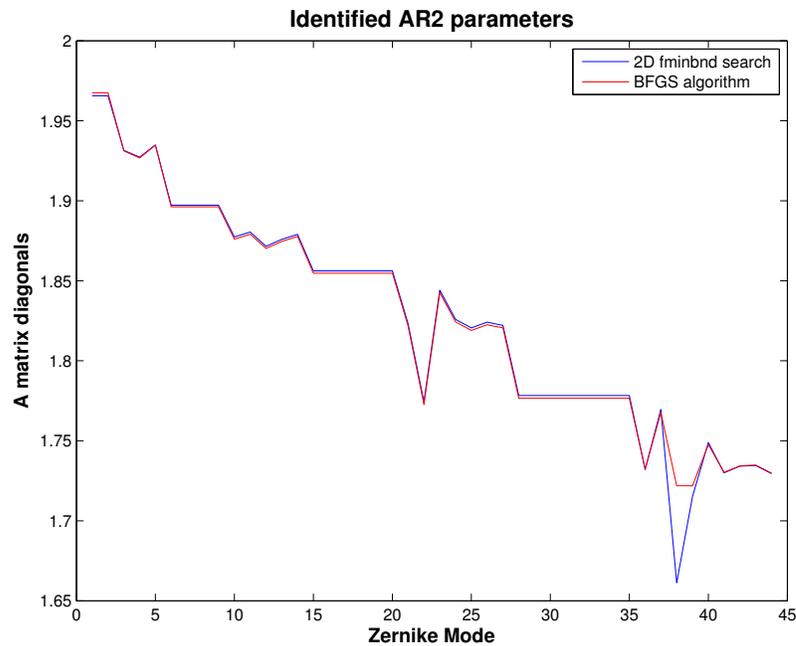


Figure 3.3: Parameter identification using different search minimization algorithms. BFGS search provides a speed up factor of more than 2 orders of magnitude in the AR2 case.

### Discretized transfer function

Once the coefficients of Eq. 3.69 have been determined, a discrete time representation of  $H(s)$  is required. The relationship between  $s$  and  $z$ ,

$$z = e^{sT_s}, \quad (3.72)$$

$$s = \frac{\log z}{T_s} \quad (3.73)$$

can be substituted, in the AR2 case, directly into Eq. 3.69,

$$H_j(z) = \frac{1}{\frac{\log^2 z}{T_s} - a_j \frac{\log z}{T_s} - b_j}, \quad (3.74)$$

however this form of  $H(z)$  does not fit well with the intended AR filtering of Eq. 3.49. A similar type of problem was addressed in [67] in which they applied the adapted transformation given in [79] which maintains the poles of the transfer function. The transformation is carried out by equating the following,

$$\frac{1}{s - p_{ji}} = \frac{T_s}{1 - e^{p_{ji}T_s} z^{-1}}, \quad (3.75)$$

where  $p_{ji}$  is the  $i^{th}$  pole of  $H_j(s)$ . Factoring the denominator of  $H_j(s)$  for the AR2 case gives the poles of the continuous transfer function in terms of the coefficients,

$$p_{j1} = \frac{-a_j + \sqrt{a_j^2 - 4b_j}}{2}, \quad (3.76)$$

$$p_{j2} = \frac{-a_j - \sqrt{a_j^2 - 4b_j}}{2} \quad (3.77)$$

Substituting the corresponding terms of Eq. 3.75 into Eq. 3.62, the discrete transfer function of mode  $j$  is written,

$$H_j(z) = \frac{T_s^2}{(1 - e^{p_{j1}T_s} z^{-1})(1 - e^{p_{j2}T_s} z^{-1})}, \quad (3.78)$$

$$= \frac{T_s^2}{1 - \underbrace{(e^{p_{j1}T_s} + e^{p_{j2}T_s})}_{z_j^a} z^{-1} + \underbrace{e^{(p_{j1}+p_{j2})T_s}}_{z_j^b} z^{-2}}. \quad (3.79)$$

To verify that the discretized model has the same response as the continuous

model, the bode plots were compared (Fig. 3.4). The coefficients of the  $z^{-1}$  and  $z^{-2}$

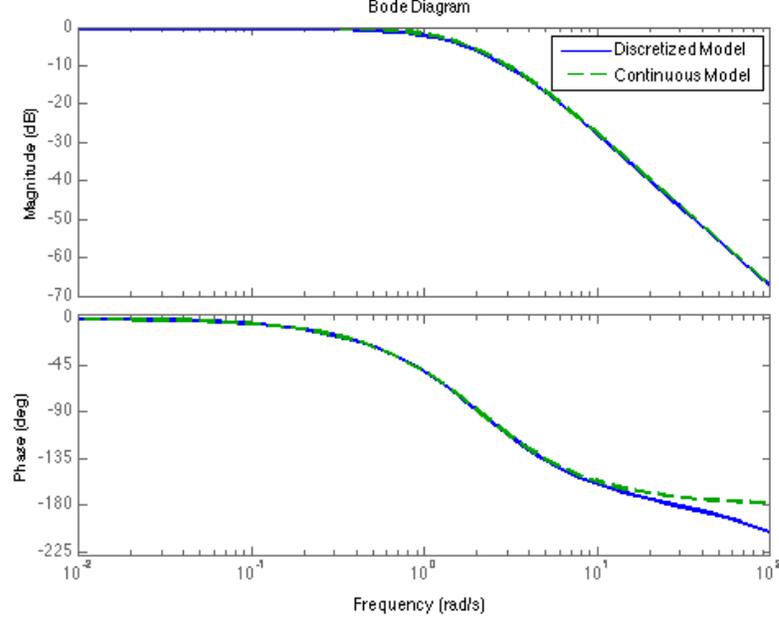


Figure 3.4: Bode plots of continuous and discrete TFs.

terms are functions of the poles of the continuous transfer function,  $z_j^a(p_1, p_2)$ ,  $z_j^b(p_1, p_2)$ , and now populate the discrete time prediction model matrices  $\mathbf{A}_{AR2}$  and  $\mathbf{B}_{AR2}$  in Eq. 3.49. For a set of Zernike modes from 1 to  $n_z$ , this second order prediction takes on the following structure for a single layer phase vector,

$$\begin{bmatrix} \hat{\varphi}_{k+1} \\ \hat{\varphi}_k \end{bmatrix} = \begin{bmatrix} z_1^a & & & z_1^b \\ & z_2^a & & z_2^b \\ & & \ddots & \ddots \\ & & & z_{n_z}^a & z_{n_z}^b \\ \hline 1 & & & & 0 \\ & 1 & & & 0 \\ & & \ddots & & \ddots \\ & & & 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{\varphi}_k \\ \hat{\varphi}_{k-1} \end{bmatrix}. \quad (3.80)$$

For a multi-layered atmosphere, a concatenation of all layer coefficients,  $[z^a(l_1), z^a(l_2), \dots]$  and  $[z^b(l_1), z^b(l_2), \dots]$ , populate the diagonals. In the AR1 case, the discrete form of the transfer function can be obtained directly by substitution from Eq. 3.75 using

directly  $a_j = p_{j1}$ ,

$$H(z) = \frac{T_s}{1 - e^{p_{j1}T_s} z^{-1}}. \quad (3.81)$$

The plots in Fig. 3.5 show the temporal auto-correlation functions from Zernike polynomials 2 to 9 using the Wiener-Khinchine theorem and the spatio-angular formulation. There is good general agreement between the two methods. Also over-plotted are the second-order continuous model and its auto-regressive discretization for  $T_s = 10 \text{ ms}$  (though any frequency can be selected), when fitting the initial  $50 \text{ ms}$ .

### Spatio-Angular first order prediction model

A more general method that complies with tomographic phase estimation is sought that fully takes into account the temporal cross-correlations of modes. Such cross-correlations are known to exist and to evolve as time elapses. Thanks to Eq. 3.40 the time-evolution of the cross-correlation between Zernike polynomials can be assessed [71]. Figure 3.6 depicts the modal cross-correlation functions for tip and focus (polynomials 2 and 4 respectively [6]) up to the 9th mode. Exploiting this knowledge is paramount to obtain the best temporal prediction possible.

Recasting the problem as a proper criterion minimization, the best linear predictor (in the MMSE sense) is the solution to the following criterion

$$\mathcal{A}_\delta^* = \arg \min_{\mathcal{A}_\delta} \langle \|\varphi(t + T_s) - \mathcal{A}_\delta \varphi(t)\|_{L_2(\Omega)}^2 \rangle, \quad (3.82)$$

Restricting the sample time,  $T_s$  to a fixed value means  $\varphi(t+T_s) \triangleq \varphi_{k+1}$  and  $\varphi(t) \triangleq \varphi_k$  which are the sampled phase vectors with  $T_s$  lag in between. The solution is found to be

$$\mathcal{A}_\delta^* \triangleq \langle \varphi_{k+1} \varphi_k^\top \rangle \langle \varphi_k \varphi_k^\top \rangle^{-1}. \quad (3.83)$$

This predictor is in what follows called the *1-step SA predictor*. It is a non-diagonal, densely populated matrix – a sign of the temporal cross-correlations between modes explored in Fig. 3.6 coming into play. The same predictor is outlined by [60] for the case of phase represented by its samples on a regular grid of points (also called the zonal representation).

Furthermore, Eq. 3.83 is a general method to generate and predict phase in a 2D plane (any wind velocity can be used) according to the Quasi-Markovian model

$$\varphi_{k+1} = \mathcal{A}_\delta^* \varphi_k + \varepsilon_k^\delta \quad (3.84)$$

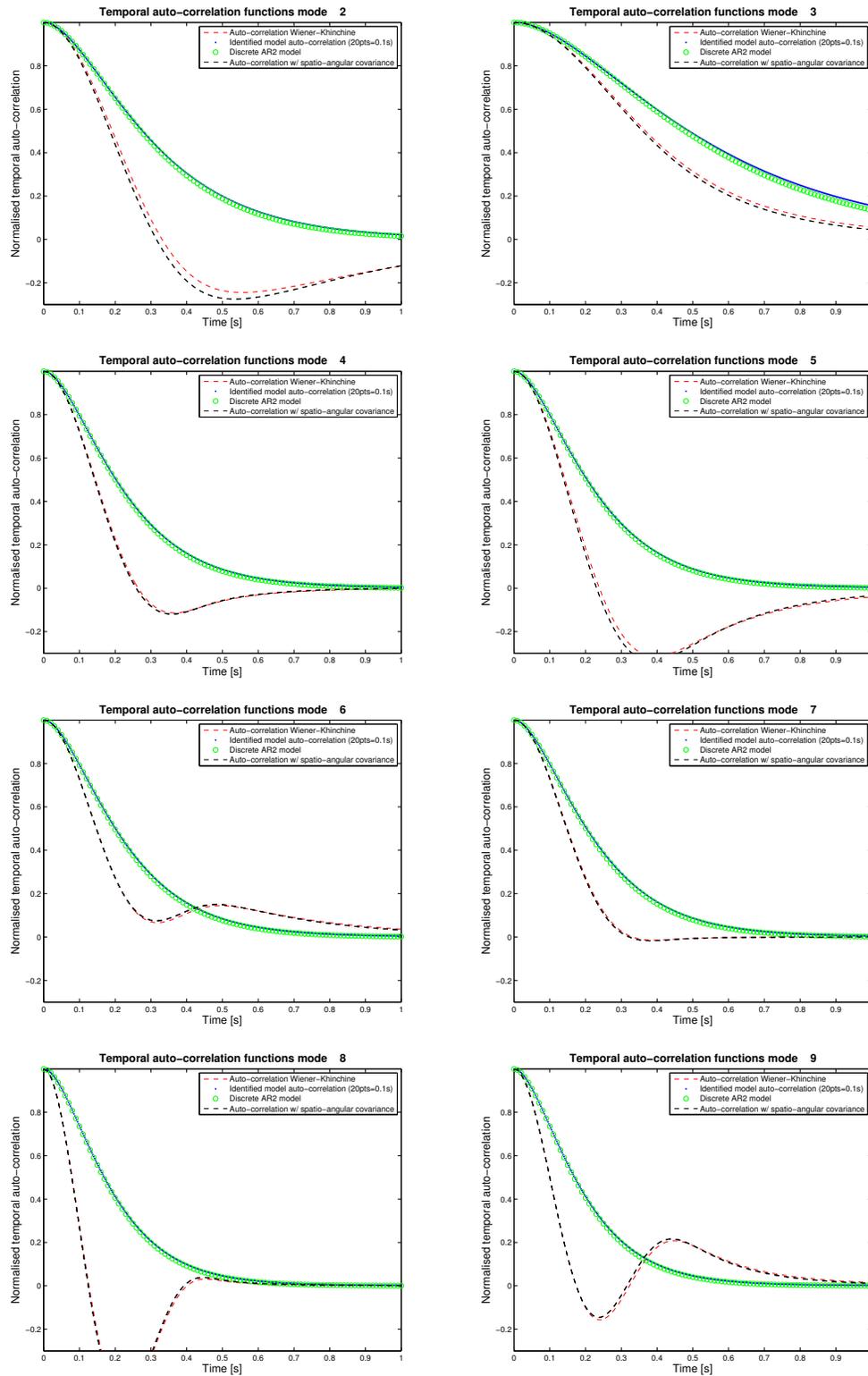


Figure 3.5: Theoretical temporal auto-correlation functions assuming frozen-flow against the 2nd-order continuous and discrete predictive models fitting the initial 50 ms of the theoretical curves.

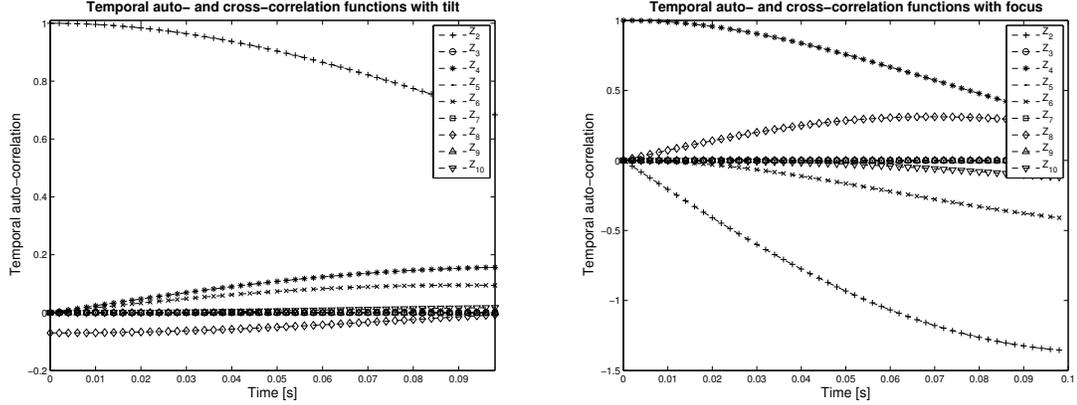


Figure 3.6: Temporal auto and cross-correlation functions for the tip and focus modes. Although at  $t = 0$  these plots show the spatial values in [6], for  $t > 0$  correlations appear and vanish as shown. One can see, for the case of focus, that although it isn't correlated to any other mode plotted for  $t = 0$ , a strong (anti) correlation appears with tilt as time elapses.

where  $\varepsilon_k^\delta$  is an excitation noise whose properties are fixed to guarantee proper turbulence statistics.

This model can either be used for generating a fully developed turbulence, or for control-oriented purposes such as a Kalman filter gain computation [61]. In either case the pair  $(\mathcal{A}_\delta^*, \Sigma_\varepsilon^\delta)$  is required. The excitation noise covariance matrix  $\Sigma_\varepsilon^\delta$  is found from imposing the output statistics to be those of a Kolmogorov or von Kármán model. Hence  $\Sigma_\varepsilon^\delta = \Sigma_\varphi - \mathcal{A}_\delta^* \Sigma_\varphi \mathcal{A}_\delta^{*\top}$ , since  $\langle \varphi_{k+1} (\varphi_{k+1})^\top \rangle = \mathcal{A}_\delta^* \Sigma_\varphi \mathcal{A}_\delta^{*\top} + \Sigma_\varepsilon^\delta$ . Note  $\Sigma_\varphi \triangleq \langle \varphi_{k+1} \varphi_{k+1}^\top \rangle \triangleq \langle \varphi_k \varphi_k^\top \rangle$ , which, due to stationarity, loses its temporal dependence.

Assémat's method for simulating infinitely long, non-stationary phase screens [80] is based on a truncated version of Eq. 3.83 when a point-wise (zonal) representation of the phase is used. However, only a sub-set of the columns of  $\mathcal{A}_\delta^*$  is used to include bounded-region correlations. An extension can be found in [81, 82].

Despite the temporal cross-correlation being taken into account in this model,  $\mathcal{A}_\delta^*$  is strongly diagonally dominated, suggesting that simpler diagonal models, *i.e.* the mode-by-mode AR models, could be potentially applied with equal results (the AR models have been extensively used in AO simulations [83]). A theoretical examination of the reduction in temporal lag error for both methods is required.

### 3.3.3 Theoretical temporal lag errors

The predictive capabilities of different models can be assessed using the predicted phase temporal structure function. Define the temporal lag error (**no** prediction) as the average phase variance taken two instants apart, *i.e.* the temporal structure function of the phase

$$D_t(\tau) = \langle |\varphi(\boldsymbol{\rho}, t) - \varphi(\boldsymbol{\rho}, t + \tau)|^2 \rangle \quad (3.85)$$

which relates to the spatial structure function of phase  $D_\varphi$  through Taylor's frozen-flow hypothesis by

$$D_t(\tau) = \langle |\varphi(\boldsymbol{\rho}, t) - \varphi(\boldsymbol{\rho} - |v|\tau, t)|^2 \rangle_\rho = D_\rho(|v|\tau) \quad (3.86)$$

with  $\langle \dots \rangle_\rho$  the ensemble average over the spatial distances  $\rho$ . Expanding the squared term and assuming stationarity gives  $2(\mathcal{C}(0) - \mathcal{C}_t(\tau)) = 2(\mathcal{C}(0) - \mathcal{C}_\rho(|v|\tau))$ , where  $\mathcal{C}_t(\cdot)$  and  $\mathcal{C}_\rho(\cdot)$  are the phase temporal and spatial covariance functions developed in Sec. 2.1.1.  $\mathcal{C}_t(\tau) = \mathcal{C}_\rho(|v|\tau)$  *i.e.* the temporal covariance function is deduced from the phase spatial covariance function evaluated at  $\rho = |\boldsymbol{\rho}| = |v|\tau$ . The phase spatial structure function is developed in Sec. 2.1.3 and given for a finite outer scale in Eq. 2.34.

It is useful to consider the general formulation for the temporal lag error that is compatible with the case of atmospheric prediction

$$\sigma^2(\tau) = \left\langle \|\varphi_k - \widehat{\varphi}_k\|_{L_2(\Omega)}^2 \right\rangle, \quad (3.87)$$

where  $\widehat{\varphi}_k$  is the phase estimate using any of the predictive models.

In the **no** prediction case the estimated phase is simply a replication of the phase at the previous time step,  $\widehat{\varphi}_k = \varphi_{k-1}$ . The temporal lag error from Eq.(3.87) becomes

$$\sigma_0^2(\tau) = \text{trace} \{ \boldsymbol{\Sigma}_\varphi - \boldsymbol{\Sigma}_\varphi^1(\tau) \} \triangleq D_t(\tau), \quad (3.88)$$

which is simply the temporal structure function of phase from Eq 3.86, with the 1-step covariance matrix  $\boldsymbol{\Sigma}_\varphi^1 = \boldsymbol{\Sigma}_\theta(\theta = |v|\tau)$

These temporal structure functions can now be expanded for the case of predicted phase using various prediction models. The theoretical temporal lag error for the modal prediction models can be computed and compared to the **no** prediction tem-

poral lag error. With the first order models (the AR1 or the SA predictor) one has directly Eq. 3.88,

$$\sigma_{\text{lag}}^2(\tau, p = 1) = \text{trace} \{ \mathbf{P}_\theta (\boldsymbol{\Sigma}_\varphi + \mathcal{A} \boldsymbol{\Sigma}_\varphi \mathcal{A}^\top - 2 \boldsymbol{\Sigma}_{1\tau} \mathcal{A}^\top) \mathbf{P}_\theta^\top \}, \quad (3.89)$$

with  $\mathcal{A} = \mathbf{A}_{\text{AR1}}$  for the AR1 model and  $\mathcal{A} = \mathcal{A}_\delta^*$  for the SA predictor model. For the AR2

$$\begin{aligned} \sigma_{\text{lag}}^2(\tau, p = 2) = \text{trace} \{ & \mathbf{P}_\theta (\boldsymbol{\Sigma}_\varphi + \mathbf{A}_{\text{AR2}} \boldsymbol{\Sigma}_\varphi \mathbf{A}_{\text{AR2}} + \mathbf{B}_{\text{AR2}} \boldsymbol{\Sigma}_\varphi \mathbf{B}_{\text{AR2}} + \\ & - 2 \mathbf{A}_{\text{AR2}} \boldsymbol{\Sigma}_{1\tau} + 2 \mathbf{A}_{\text{AR2}} \boldsymbol{\Sigma}_{1\tau} \mathbf{B}_{\text{AR2}}^\top - 2 \mathbf{B}_{\text{AR2}} \boldsymbol{\Sigma}_{2\tau}) \mathbf{P}_\theta^\top \} \end{aligned} \quad (3.90)$$

with  $\boldsymbol{\Sigma}_{1\tau} = \boldsymbol{\Sigma}(\rho = v\tau)$  and  $\boldsymbol{\Sigma}_{2\tau} = \boldsymbol{\Sigma}(\rho = 2v\tau)$ .

Likewise for the AR3

$$\begin{aligned} \sigma_{\text{lag}}^2(\tau, p = 3) = \text{trace} \{ & \mathbf{P}_\theta (\boldsymbol{\Sigma}_\varphi + \mathbf{A}_{\text{AR3}} \boldsymbol{\Sigma}_\varphi \mathbf{A}_{\text{AR3}} + \mathbf{B}_{\text{AR3}} \boldsymbol{\Sigma}_\varphi \mathbf{B}_{\text{AR3}} + \\ & + \mathbf{C} \boldsymbol{\Sigma}_\varphi \mathbf{C}_{\text{AR3}} - 2 \mathbf{A}_{\text{AR3}} \boldsymbol{\Sigma}_{1\tau} - 2 \mathbf{B}_{\text{AR3}} \boldsymbol{\Sigma}_{2\tau} - 2 \mathbf{C}_{\text{AR3}} \boldsymbol{\Sigma}_{3\tau} + \\ & + 2 \mathbf{A}_{\text{AR3}} \boldsymbol{\Sigma}_{1\tau} \mathbf{B}_{\text{AR3}}^\top + 2 \mathbf{B}_{\text{AR3}} \boldsymbol{\Sigma}_{1\tau} \mathbf{C}_{\text{AR3}}^\top + 2 \mathbf{A}_{\text{AR3}} \boldsymbol{\Sigma}_{2\tau} \mathbf{C}_{\text{AR3}}^\top) \mathbf{P}_\theta^\top \} \end{aligned} \quad (3.91)$$

with  $\boldsymbol{\Sigma}_{3\tau} = \boldsymbol{\Sigma}(\rho = 3v\tau)$ . The matrices  $\boldsymbol{\Sigma}_{p\tau}$  are computed for a fully-developed turbulence following Eq. 3.40 and thus differ from those associated with any specific model.

The theoretical error introduced due to pure system lag of increasing length is shown in Fig. 3.7 along with the performance of all the predictive models presented above.

- The 1-step SA predictor provides the best performance (as expected) for large lags above  $\sim 5$  ms, which stems from its optimality (it minimizes the prediction error variance in Eq. 3.82) and cross-mode prediction. However, the predictor is only a truncated version: in practice a finite number of modes is to be used. When more modes are added in, the 1-step SA predictor beats the AR2-3 models for lower lags,
- For lags below  $\sim 5$  ms AR models of orders 2 and 3 slightly outperform the 1-step SA predictor estimating the same number of modes since the short term decorrelation of the AR models is quite similar to that of the turbulence.
- the AR2 is a sufficient model-order above which the performance gains are

little and not worth the increased complexity for both parameter identification or real-time processing; an AR2 model presents a temporal PSD  $\propto \nu^{-4}$  which is a good approximation to the actual phase whose spectrum is  $\propto \nu^{-11/3}$ , with  $\nu = |\boldsymbol{\nu}|$  the temporal frequency vector modulus,

- although the AR1 has been successfully used in control-oriented models for Kalman filtering [83, 66, 84], it performs quite poorly to predict phase and is patently incapable of providing any improvement over the no prediction case. Simulations show that when the off-diagonal values of the 1-step SA predictor are nulled out, the predictive capability degrades to that of a diagonal AR1, suggesting that it is this feature that plays the most important role in the predictive process.

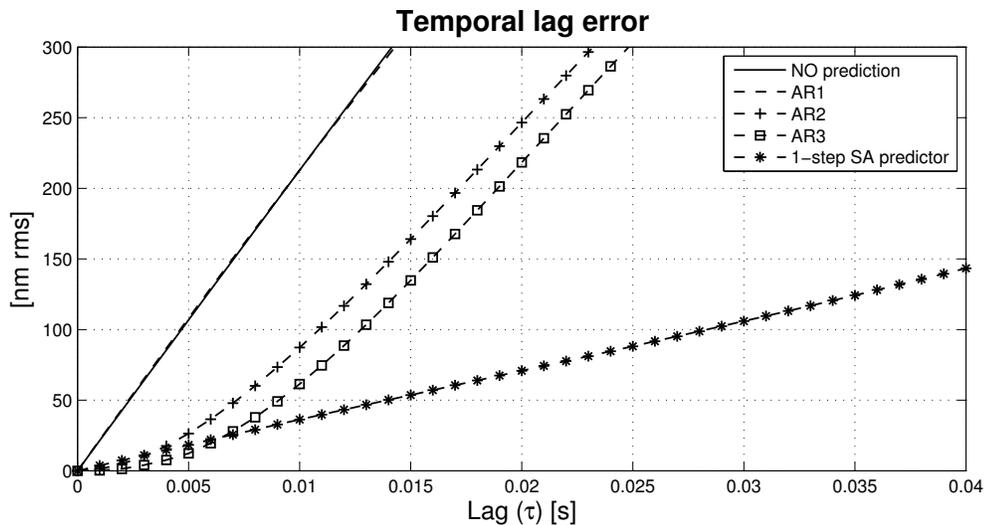


Figure 3.7: Comparison of temporal lag errors on an equivalent single layer atmosphere see Table 4.2 for further parameters.

Approaching this from a different point of view, it can be seen that potential gains in performance can be traded for an increase in limiting magnitude. In the presence of temporal prediction, the lag error  $\sigma_{\text{lag}}^2$  decreases and one can thus achieve the same performance as the static case when tolerating more measurement noise from fainter sources. Increased measurement noise, resulting in increased noise propagation  $\sigma_{np}^2$  (averaged over the field) for the same aggregate wave-front error, is defined in the

first order as

$$\sigma_{\text{total}}^2 = \underbrace{\sigma_{\text{lag}}^2 + \sigma_{np}^2}_{\sigma^2} + \text{other terms} \quad (3.92)$$

where 'other terms' relates to errors that are independent from the system lag (to a first degree approximation). The latter is computed from Eq. 3.87 whereas the noise propagation error is

$$\sigma_{np}^2 = \text{trace} \{ \mathbf{R} \boldsymbol{\Sigma}_\eta \mathbf{R}^\top \}, \quad (3.93)$$

where the noise covariance matrix  $\boldsymbol{\Sigma}_\eta = \sigma_\eta^2 \mathbf{I}$  is assumed diagonal with  $\sigma_\eta^2$  the measurement noise variance on each WFS sub-aperture. The latter is a decreasing function of the star brightness, *i.e.* photon-noise increases for dimmer guide-stars. The propagated noise  $\sigma_{np}^2$  relates to star magnitude by standard centroiding error functions [85].

The AO system bandwidth is chosen to minimize the error  $\sigma^2 = \sigma_{\text{lag}}^2 + \sigma_{np}^2$ . Thus one has

$$\sigma^2 = \sigma_{np}^2(m_v, \tau) + \sigma_{\text{lag}}^2(\tau, p = 0) = \sigma_{np}^2(m'_v, \tau) + \sigma_{\text{lag}}^2(\tau, p > 0) \quad (3.94)$$

where  $\sigma_{\text{lag}}^2(\tau, p = 0) \geq \sigma_{\text{lag}}^2(\tau, p > 0)$  and  $\sigma_{np}^2(m_v, \tau) \leq \sigma_{np}^2(m'_v, \tau)$ , with  $m'_v - m_v$  the limiting magnitude increase.

This rather coarse approximation will give some insight into the potential magnitude increase factor. Figure 3.8 plots the result  $m'_v - m_v$  as a function of lag. The noise propagation coefficient used was computed to be  $\sigma_{np}^2/\sigma_\eta^2 \approx 0.5$ , following an extension to tomography of the analytical derivation in [86].

As expected, the minima in  $\sigma^2$  (in blue, ordinate on the right) are achieved for longer integration times as the GS are fainter. These minima are indicated by vertical dotted lines, for which red circles indicate the increased limiting magnitudes when the sources vary from magnitudes 13 to 17. A consistent value around half a magnitude increase is obtained. Only a full end-to-end Monte Carlo simulation can inform about the effective magnitude increase gain, considering the actual EE and SR figures of merit; this numerical simulation is carried out in Chapter 4.

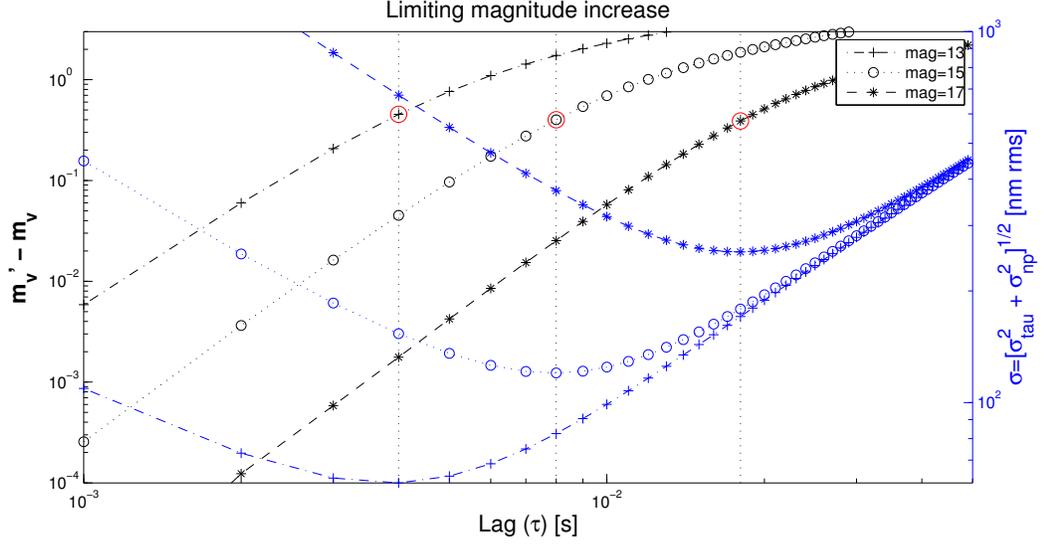


Figure 3.8: Temporal lag error is traded by increased noise propagation through the wave-front reconstruction. Black markers: increase in limiting magnitude; blue markers: combined temporal plus noise propagation error; red circles indicate the increased limiting magnitudes for the minima of  $\sigma$ .

### 3.3.4 Linear Quadratic Gaussian controller with modal prediction models

Based on the analysis of the various prediction models, the LQG controller formulated in Sec.4.3.3 was developed using two state transition models: a second order AR model, and a first order SA model, both applying an explicit layer-by-layer prediction. In the case of the AR2 model, the state  $\mathbf{x}_k$  contains, at any given time, two instances of the phase estimate vector as the phase estimate at time  $k+1$  is a linear combination of phase from  $k$  and  $k-1$ ,

$$\hat{\boldsymbol{\varphi}}_{k+1} = \mathbf{A}_{\text{AR2}}\hat{\boldsymbol{\varphi}}_k + \mathbf{B}_{\text{AR2}}\hat{\boldsymbol{\varphi}}_{k-1}. \quad (3.95)$$

Prediction matrices  $\mathbf{A}_{\text{AR2}}$  and  $\mathbf{B}_{\text{AR2}}$  are computed for each layer following the method laid out in Sec. 3.3.2.

In the case of any first order prediction model, only the current measurement (corresponding to a specific time in the past) is used and no additional data is required. A single state LQG controller can therefore be implemented, leading to a significant reduction in the size of the state matrix. Given the promising performance projected to be achieved by the first order SA predictor, a single state LQG controller can be

implemented as follows if a one-step delay is assumed,

$$\begin{aligned}
\hat{\mathbf{s}}_{k|k-1}^{(\alpha)} &= \mathbf{\Gamma} \mathbf{P}_\alpha \hat{\boldsymbol{\varphi}}_{k|k-1} \\
\hat{\boldsymbol{\varphi}}_{k|k} &= \hat{\boldsymbol{\varphi}}_{k|k-1} + \mathcal{M}_\infty (\mathbf{s}_k - \hat{\mathbf{s}}_{k|k-1}) \\
\hat{\boldsymbol{\varphi}}_{k+1|k} &= \mathcal{A}_\delta^* \hat{\boldsymbol{\varphi}}_{k|k} \\
\mathbf{u}_k &= (\mathbf{I} \mathbf{M})^\dagger \mathbf{\Gamma} \mathbf{P}_\beta \hat{\boldsymbol{\varphi}}_{k+1|k}.
\end{aligned} \tag{3.96}$$

The AR2 state transition matrices,  $\mathbf{A}_{\text{AR2}}$  and  $\mathbf{B}_{\text{AR2}}$ , are block diagonal matrices with each block, a diagonal matrix of coefficients corresponding to the AR2 model for a given layer in the atmosphere whereas the prediction matrix,  $\mathcal{A}_\delta^*$  is a block diagonal matrix where each block is a full matrix corresponding to the SA prediction for a given layer. So the dimensions of the state vector are smaller for the SA case, but the state transition is more complex. In both cases, the dimensions are still driven by the number of estimated Zernike modes and the number of layers in the atmosphere model. The Kalman gain,  $\mathcal{M}_\infty$  generates, by construction, an estimate of the phase in each layer. Its dimensions are driven by the number of Zernike modes, the number of layers and the number of WFS slope measurements. One can imagine that, if a large number of layers are being estimated, these matrices could become large and cumbersome.

The additional asynchronous prediction step can also be carried out before projection onto the DM commands by applying  $\mathcal{A}_\Delta$  computed from any arbitrary prediction model, although we used the first order SA model in all cases.

### 3.4 Zonal Reconstructors

A parallel set of tomographic reconstructors following the same progression from static to predictive to dynamic can be developed with a point-wise representation of the phase. In this zonal approach, the SH-WFS measurement model remains unchanged,

$$\mathbf{s} = \mathbf{\Gamma} \boldsymbol{\varphi} + \boldsymbol{\eta}, \tag{3.97}$$

however  $\mathbf{\Gamma}$  is now a linear transformation from discrete phase points, representing specific spatial locations, to SH-WFS slopes. It is a two-dimensional concatenation of

both regular (Simpson weights) stencils and modified stencils to account for physical edge effects where the pupil of the system leads to partially illuminated WFS subapertures. Each of the modified stencils are computed individually according to the specific nature of the partial illumination - i.e. how many grid points are cut off by the pupil, and the configuration of the remaining points. The stencil grid up samples the WFS lenslet grid by a factor of 2, resulting in a  $3 \times 3$  stencil overlaid on each subaperture whose corners make a  $2 \times 2$  grid. The regular stencil has the Simpson weights for a fully illuminated subaperture,

$$\begin{bmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}. \quad (3.98)$$

The stencils expand the phase as a bilinear spline and compute the phase in each of the quadrants for each subaperture. The stencil can be generalized for any partially illuminated subaperture: according to the work done by [87] for the TMT facility AO system, NFIRAOS,

$$\begin{bmatrix} s_{1,1} & s_{1,2} & s_{1,3} \\ s_{2,1} & s_{2,2} & s_{2,3} \\ s_{3,1} & s_{3,2} & s_{3,3} \end{bmatrix} \quad (3.99)$$

The values  $s_{i,j}$  are computed as follows,

$$\begin{aligned} s_{1,1} &= -\frac{2(a_{1,1}+a_{1,1})+a_{2,1}+a_{2,2}}{3(a_{1,1}+a_{1,2}+a_{2,1}+a_{2,2})} & s_{1,3} &= \frac{2(a_{1,2}+a_{1,3})+a_{2,2}+a_{2,3}}{3(a_{1,2}+a_{1,3}+a_{2,2}+a_{2,3})} \\ s_{3,1} &= -\frac{a_{2,1}+a_{2,2}+2(a_{3,1}+a_{3,2})}{3(a_{2,1}+a_{2,2}+a_{3,1}+a_{3,2})} & s_{3,3} &= \frac{a_{2,2}+a_{2,3}+2(a_{3,2}+a_{3,3})}{3(a_{2,2}+a_{2,3}+a_{3,2}+a_{3,3})} \\ s_{1,2} &= -(s_{1,1} + s_{1,3}) & s_{3,2} &= -(s_{3,1} + s_{3,3}) \\ s_{2,1} &= -(s_{1,1} + 2s_{3,1} + 2) & s_{2,3} &= -(s_{1,3} + s_{3,3}) + 2 \\ s_{2,2} &= -(s_{2,1} + s_{2,3}). \end{aligned} \quad (3.100)$$

The values  $a_{i,j}$  are the weights given in Matrix 3.98.

### 3.4.1 Zonal static MMSE reconstructors

#### Spatio-Angular MMSE

In this zonal approach, to compute the SA covariance matrix, the separation vector,  $\mathbf{r}$  is defined as a function of the altitude of the layer,  $h_l$ , and the coordinates of the subapertures,  $\boldsymbol{\rho}$ . Computing the covariance between lenslets on different WFSs

requires the position in the pupil of subaperture  $i$  on WFS  $n$ :  $(\boldsymbol{\rho}_{i,n})$ , projected into layer  $l$  at  $h_l$  in the direction of that WFS,  $(\boldsymbol{\theta}_n)$ . The coordinates of the the subaperture in any layer are,

$$[(\rho_{x_{i,n}} + \theta_{x_n} h_l), (\rho_{y_{i,n}} + \theta_{y_n} h_l)], \quad (3.101)$$

and the separation vector between subaperture  $i$  on WFS  $n$  and subaperture  $j$  on WFS  $m$  at layer  $l$  is,

$$\mathbf{r}_{i,j,l} = (\boldsymbol{\rho}_{i,n} - \boldsymbol{\rho}_{j,m}) + h_l(\boldsymbol{\theta}_n - \boldsymbol{\theta}_m), \quad (3.102)$$

where the first term of the right hand side of the equation is the particular separation between the subapertures in the pupil and the second term is the global WFS separation vector. From Eq. 2.13, the phase spatial covariance function for finite outer scale, the SA covariance matrix between WFSs  $n$  and  $m$  can be computed for a given  $C_n^2$  profile with  $N_l$  layers for every baseline,  $\mathbf{r}$ ,

$$\boldsymbol{\Sigma}(\mathbf{r}, r_0, L_0)_{\phi}^{n,m} = \sum_{l=1}^{N_l} f r_{0_l} \boldsymbol{\Sigma}_{\phi}(\mathbf{r}_l, r_0, L_0), \quad (3.103)$$

where  $f r_{0_l}$  is the fractional Fried parameter  $r_0$ ,  $\sum_l f r_{0_l} = 1$ ,  $L_0$  is the outer scale and  $\mathbf{r}_l$  is a concatenation of all the baselines  $i, j$ . To simplify notation, this expression is referred to as  $\boldsymbol{\Sigma}_{n,m}$ .

To be able to use the phase covariance matrices with the input slope measurements, use the phase to slopes model,  $\boldsymbol{\Gamma}$ ,

$$\hat{\boldsymbol{\phi}}_{\beta} = \boldsymbol{\Sigma}_{\beta,\alpha} \boldsymbol{\Gamma}^{\top} (\boldsymbol{\Gamma} \boldsymbol{\Sigma}_{\alpha,\alpha} \boldsymbol{\Gamma}^{\top} + \boldsymbol{\Sigma}_{\eta})^{-1} \mathbf{s}_{\alpha}, \quad (3.104)$$

which is equivalent to Eq. 3.16. Once again, the phase is mapped back onto slopes in the  $\boldsymbol{\beta}$  directions due to the DM calibration available from the system,

$$\hat{\mathbf{s}}_{\beta} = \boldsymbol{\Gamma} \hat{\boldsymbol{\phi}}_{\beta} = \boldsymbol{\Gamma} \boldsymbol{\Sigma}_{\beta,\alpha} \boldsymbol{\Gamma}^{\top} (\boldsymbol{\Gamma} \boldsymbol{\Sigma}_{\alpha,\alpha} \boldsymbol{\Gamma}^{\top} + \boldsymbol{\Sigma}_{\eta})^{-1} \mathbf{s}_{\alpha}. \quad (3.105)$$

### Zonal explicit layered MMSE

In the case of the explicit layered MMSE reconstructor in zonal space, the phase covariance is computed in each layer of the meta-pupil, that is, the number of phase points increases with layer height to generate  $\langle \boldsymbol{\varphi} \boldsymbol{\varphi}^{\top} \rangle$  in each layer according to the

given sampling and independent of direction. The propagation matrix  $\mathbf{P}_\theta$  introduced in Sec. 3.2.1 is formulated as the zonal propagation matrix,  $\mathbf{H}_\theta$ , which crops the covariance in each layer in each direction and propagates it into the pupil.

$$\hat{\mathbf{s}}_\beta = \mathbf{\Gamma} \mathbf{H}_\beta \langle \varphi \varphi^\top \rangle \mathbf{H}_\alpha^\top \mathbf{\Gamma}^\top (\mathbf{\Gamma} \mathbf{H}_\alpha \langle \varphi \varphi^\top \rangle \mathbf{H}_\alpha^\top \mathbf{\Gamma}^\top + \mathbf{\Sigma}_\eta)^{-1}, \quad (3.106)$$

where the propagation matrices,  $\mathbf{H}_\alpha, \mathbf{H}_\beta$ , are computed by a bilinear spline interpolation between the phase points in the ground grid,  $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$  and the phase points in the grid in the direction defined by the star coordinates in zenith and azimuth angle,  $(\boldsymbol{\xi}, \boldsymbol{\theta})$ , and the height,  $h_i$ , of each atmospheric layer,  $\mathbf{z}_i = (\mathbf{x}_i, \mathbf{y}_i)$ ,

$$\begin{aligned} \mathbf{x}_i &= h_i \tan \boldsymbol{\xi} \cos \boldsymbol{\theta}, \\ \mathbf{y}_i &= h_i \tan \boldsymbol{\xi} \sin \boldsymbol{\theta}. \end{aligned} \quad (3.107)$$

A schematic of the coordinate system is shown in Fig. 3.9a and a plot of the actual phase-point locations relative to each other for a particular case is shown in Fig. 3.9b.

In the case of astronomical objects (NGSs or science targets), which are located at infinity, the grid spacing,  $d_0$ , is the same in both ground and layer grids and the offset of the phase points in the layer relative to the ground are defined as,

$$\begin{aligned} \mathbf{u} &= \frac{\mathbf{x}_i - \mathbf{x}_0}{d_0} \\ \mathbf{v} &= \frac{\mathbf{y}_i - \mathbf{y}_0}{d_0}. \end{aligned} \quad (3.108)$$

In the LGS case, the grid is scaled according to the altitude of the layer and the distance of the LGS from the telescope to account for the cone effect. Using these coordinate grids,  $\mathbf{H}(\boldsymbol{\xi}, \boldsymbol{\theta})$  executes a two dimensional linear interpolation,

$$\mathbf{H} = \begin{cases} (1 - |\mathbf{u}|)(1 - |\mathbf{v}|), & |\mathbf{u}|, |\mathbf{v}| < 1 \\ 0, & \textit{elsewhere} \end{cases} \quad (3.109)$$

### 3.4.2 Predictive Spatio-Angular MMSE

Applying the Taylor frozen flow hypothesis [77] to equate spatial displacement within the pupil to temporal delay at a fixed position in the pupil, a single step pupil-plane predictor can be developed by leading out the computation in the direction of the wind profile. As a result, the reconstructor given in Eq. 3.105 is computed with a

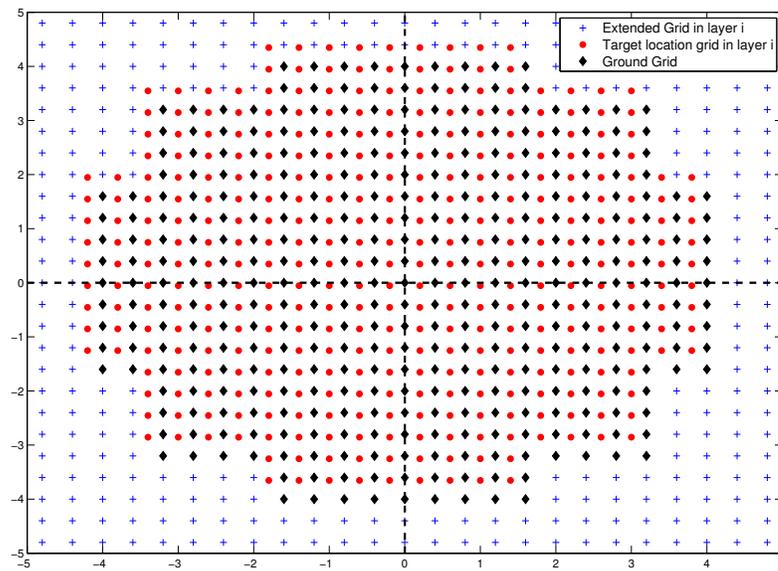
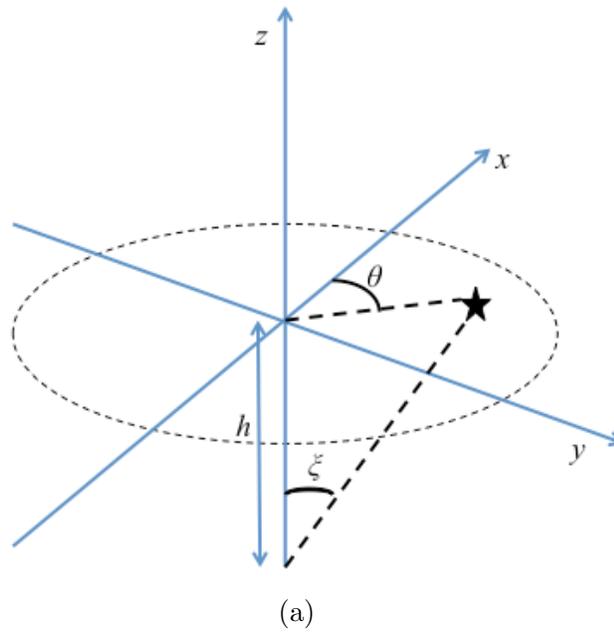


Figure 3.9: Top: Coordinate system used to define the coordinates of phase point grid centre in an atmospheric layer at altitude  $h$ . Bottom: Phase point locations: The ground grid (black diamonds) is expanded to fill the meta-pupil at layer  $i$  (blue crosses) and the target phase point locations (red dots) must be interpolated onto the ground grid locations.

modified direction vector. To compute the predictive covariance between the OL-WFS directions,  $\boldsymbol{\alpha}_n$  and the science object directions,  $\boldsymbol{\beta}_m$ , shift the global separation vector in each layer according to the wind velocity,  $\boldsymbol{v}$ , and the sample time,  $T_s$ ,

$$\boldsymbol{\delta} = \boldsymbol{v}T_s \quad (3.110)$$

$$\boldsymbol{r}_{i,j,l}^{m,n} = (\boldsymbol{\rho}_{i,n} - \boldsymbol{\rho}_{j,m}) + h_l[\boldsymbol{\alpha}_n - (\boldsymbol{\beta}_m + \boldsymbol{\delta})]. \quad (3.111)$$

Following Eq. 3.104 above, the predicted phase in direction  $\boldsymbol{\beta}$  can be expressed as

$$\phi_{k+1}^{(\boldsymbol{\beta})} = \phi_k^{(\boldsymbol{\beta}+\boldsymbol{\delta})} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}+\boldsymbol{\delta},\boldsymbol{\alpha}}\boldsymbol{\Gamma}^\top(\boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}\boldsymbol{\Gamma}^\top + \boldsymbol{\Sigma}_\eta)^{-1}\boldsymbol{s}_\alpha(k). \quad (3.112)$$

The notation  $\boldsymbol{\Sigma}_{\boldsymbol{\theta}+\boldsymbol{\delta}}$  used here, and in following sections, is representative of a covariance function adjusted by a small angle that is a function of  $\boldsymbol{\delta}$ . As in the modal SA prediction in Sec. 3.3.2, this reconstructor will have the same dimensions as the static reconstructor and adds no computational complexity to the real-time path.

This computation is effectively two steps in one, combining spatial and temporal estimation into a single step. This is fine for the stand-alone prediction, however the two steps must necessarily be split up in the zonal SA LQG controller introduced in the next section. It can be verified that doing the two steps separately is equivalent to the all-at-once reconstructor:

*Proof. All-at-once vs. two-steps scheme*

Given the properties of the conditional expectation, namely

$$E\{\Phi|S = s\} = E\{\Phi|E\{Y|S = s\}\} \quad (3.113)$$

the all-at-once and the split spatial plus temporal estimations should be the same.

Take the single step calculation

$$\widehat{\phi}_{k+1}^{(\boldsymbol{\beta})} = E\{\phi_{k+1}^{(\boldsymbol{\beta})}|\phi_k^{(\boldsymbol{\alpha})}\} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}+\boldsymbol{\delta},\boldsymbol{\alpha}}\boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}^{-1}\phi_k^{(\boldsymbol{\alpha})}, \quad (3.114)$$

and compare it to the two-step calculation which reads

$$\widehat{\phi}_{k+1}^{(\boldsymbol{\beta})} = E\{\phi_{k+1}^{(\boldsymbol{\beta})}|E\{\phi_k^{(\boldsymbol{\beta})}|\phi_k^{(\boldsymbol{\alpha})}\}\} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}+\boldsymbol{\delta},\boldsymbol{\beta}}\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\beta}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\alpha}}\boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}^{-1}\phi_k^{(\boldsymbol{\alpha})}. \quad (3.115)$$

The SA prediction model given in Eq. 3.83 can be recast as a spatial transformation

model and defined in terms of the pupil-plane phase,

$$\mathcal{A}_\delta^* = \langle \phi_\theta \phi_\alpha^\top \rangle \langle \phi_\alpha \phi_\alpha^\top \rangle^{-1}. \quad (3.116)$$

Take now  $\phi_\beta = \mathcal{A}_\delta^* \phi_\alpha$ . Thus

$$\begin{aligned} \widehat{\phi}_{k+1}^{(\beta)} &= E\{\phi_{k+1}^{(\beta)} | E\{\mathcal{A}_\delta^* \phi_k^{(\alpha)} | \phi_k^{(\alpha)}\}\} \\ &= \langle \phi_{\beta+\delta} (\mathcal{A}_\delta^* \phi_\alpha)^\top \rangle \langle (\mathcal{A}_\delta^* \phi_\alpha) (\mathcal{A}_\delta^* \phi_\alpha)^\top \rangle^{-1} \langle \mathcal{A}_\delta^* \phi_\alpha \phi_\alpha^\top \rangle \langle \phi_\alpha \phi_\alpha^\top \rangle^{-1} \phi_k^{(\alpha)} \\ &= \Sigma_{\beta+\delta, \alpha} \mathcal{A}_\delta^* (\mathcal{A}_\delta^* \Sigma_{\alpha, \alpha} \mathcal{A}_\delta^{*\top})^{-1} \mathcal{A}_\delta^* \Sigma_{\alpha, \alpha} \Sigma_{\alpha, \alpha}^{-1} \phi_k^{(\alpha)} \\ &= \Sigma_{\beta+\delta, \alpha} \Sigma_{\alpha, \alpha}^{-1} \phi_k^{(\alpha)}. \end{aligned} \quad (3.117)$$

It can also be shown numerically that  $\Sigma_{\beta+\delta, \alpha} = \Sigma_{\beta+\delta, \beta} \Sigma_{\beta, \alpha}^{-1} \Sigma_{\beta, \alpha}$ .

□

Equation 3.116 naturally gives rise to a time evolution model of the first order. The state noise covariance matrix can be likewise computed from Eq. 3.47 for the zonal case. It turns out that the method presented here generalizes that presented in [80] to fractional pixel shifts in 2-dimensional directions. Since we are concerned with phase prediction, the method works out to shift the phase screen in the appropriate wind direction with shifts given by  $|v_l| \cos(\theta_l)$  and  $|v_l| \sin(\theta_l)$  for layer  $l$ . Furthermore, the new turbulence that enters in the telescope aperture is estimated using spatial correlations with all the points in the aperture, not just a few columns of spatially close-enough turbulence as is the case for phase-screen generation in [80]. Both methods can finally be seen as a moving aperture over the phase screen where the newly seen turbulence term is generated on-the-fly. The time however is reversed in both methods.

It is important to note that this model is specific for frozen-flow, which is but part of the actual atmospheric disturbances, creating a robustness problem. A finely tuned model for frozen-flow will generate the best performance in the case of a match to the real atmosphere, but will, in all likelihood, perform more poorly for the general case. The numerical simulations and laboratory generated turbulence in this dissertation use translating phase-screens and thus do not allow to fully investigate this claim.

### 3.4.3 Bi-linear spline interpolation model

Figure 3.10 depicts an analytical covariance matrix and a thresholded  $\mathcal{A}_\delta^*$  from Eq. 3.116. It is apparent that the optimal one-step estimator is largely dominated by a diagonal term that simply shifts the phase in the appropriate direction. Note however that points outside the pupil entering the telescope (upper band in Fig.3.10-right) are computed using their respective covariance function with points inside the pupil. We thus expect the bilinear interpolation to be a reasonable approximation to the near-Markov model but with degraded performance.

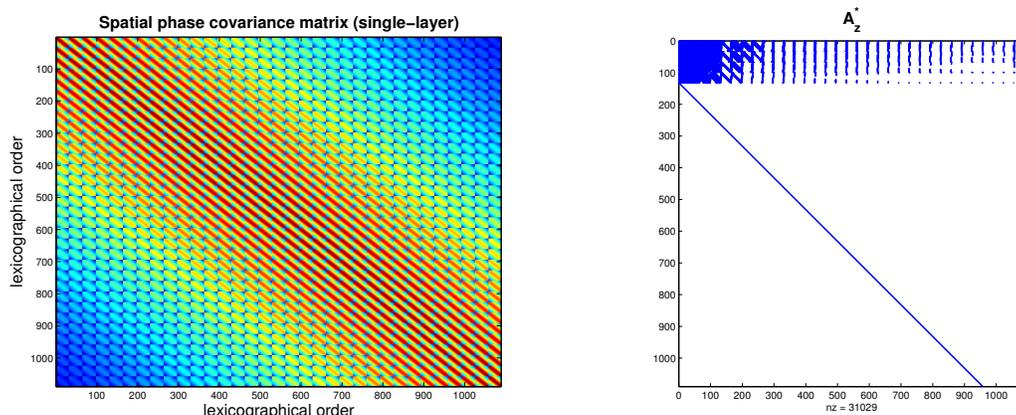


Figure 3.10: Left: Spatial phase covariance matrix (single-layer, squared pupil, sampled with 33 points across). Right: Matrix  $\mathcal{A}_\delta^*$  for an horizontal shift of 0.5m. Sampling 0.25m. The optimal one-step estimator is largely dominated by a diagonal term that simply shifts the phase in the appropriate direction. The same stands when the shifts are not integer factors of the spatial sampling. In that case,  $\mathcal{A}_\delta^*$  becomes more populated with up to four main diagonals that interpolate every single point based on correlations with neighbouring 4 points. Matrix shown for entries  $> 10^{-4}$ .

### 3.4.4 Zonal Linear Quadratic Gaussian controllers

#### Pupil-plane LQG with Spatio-Angular prediction model

The SA formulation can easily be extended to a state-space model providing a full dynamic WF reconstruction using the LQG framework. Unlike previous work [88] we resort to a full-SA LQG which has several computational advantages to add to those stated above, in particular a reduced number of states, admitting a minimal representation with a single WF instance at a given time-step.

In the pupil-plane LQG, the spatio-temporally optimal estimate of the phase in the pupil at time  $k + 1$  *in the direction of the GSs* is made,

$$\begin{aligned}\widehat{\mathbf{s}}_{k|k-1} &= \mathbf{\Gamma} \widehat{\boldsymbol{\phi}}_{k|k-1}^{(\alpha)} \\ \widehat{\boldsymbol{\phi}}_{k|k}^{(\alpha)} &= \widehat{\boldsymbol{\phi}}_{k|k-1}^{(\alpha)} + \mathcal{M}_\infty(\mathbf{s}_\alpha - \widehat{\mathbf{s}}_{k|k-1}) \\ \widehat{\boldsymbol{\phi}}_{k+1|k}^{(\alpha)} &= \mathcal{A}_\delta^* \widehat{\boldsymbol{\phi}}_{k|k}^{(\alpha)}.\end{aligned}\tag{3.118}$$

The predictive model,  $\mathcal{A}_\delta^*$  is the SA model developed for zonal space in Sec. 3.4.2, except that instead of leading off the position vector in the science direction, Eq. 3.111 is rewritten as,

$$\mathbf{r} = (\boldsymbol{\rho}_{i,n} - \boldsymbol{\rho}_{j,m}) + h_l[\boldsymbol{\alpha}_n - (\boldsymbol{\alpha}_m + \boldsymbol{\delta})],\tag{3.119}$$

for  $n, m = 1 : N_{GS}$  including the cases where  $n = m$ , with  $N_{GS}$  the total number of GSs. From this lenslet position vector, the prediction matrix,  $\mathcal{A}_\delta^*$  can be defined as,

$$\mathcal{A}_\delta^* = \boldsymbol{\Sigma}_{\boldsymbol{\alpha}+\boldsymbol{\delta},\boldsymbol{\alpha}} \boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}^{-1}.\tag{3.120}$$

It can be seen in Eq. 3.118 that the result of the controller is the estimate of the predicted phase in the GS directions. The predicted phase in the science directions are then computed from  $\widehat{\boldsymbol{\phi}}_{k+1|k}^{(\alpha)}$  using the static SA reconstructor,

$$\widehat{\boldsymbol{\phi}}_{k+1|k}^{(\beta)} = \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\alpha}} \boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}^{-1} \widehat{\boldsymbol{\phi}}_{k+1|k}^{(\alpha)}.\tag{3.121}$$

Thus executing the two-step temporal then spatial scheme discussed in Sec. 3.4.2. Once again, this algorithm is expressed assuming integer frame delays. A single frame delay can be assumed in generating the estimated phase,  $\widehat{\boldsymbol{\phi}}_{k+1|k}^{(\alpha)}$  and, subsequently, the extrapolation to the science direction given in Eq. 3.121 can be replaced by,

$$\widehat{\boldsymbol{\phi}}_{k+1+\Delta|k}^{(\beta)} = \boldsymbol{\Sigma}_{\boldsymbol{\beta}+\boldsymbol{\Delta},\boldsymbol{\alpha}} \boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}^{-1} \widehat{\boldsymbol{\phi}}_{k+1|k}^{(\alpha)}.\tag{3.122}$$

### Explicit LQG in phase space

In the explicit layer algorithm, the state vector may contain several instances of the phase vector which itself contains a vector for each layer in the atmosphere model,  $\boldsymbol{\varphi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_{N_l}]^\top$ . The dimensions of various matrices therefore scale with

the number of layers. Additional matrix multiplications which propagate the phase between the layers and the pupil ( $\mathbf{H}_\alpha$ ) must also be carried out at each time step. Assuming a single state is used, the real-time steps of the explicit layered LQG in zonal space are as follows,

$$\begin{aligned}
\widehat{\mathbf{s}}_{k|k-1} &= \mathbf{\Gamma} \mathbf{H}_\alpha \widehat{\boldsymbol{\varphi}}_{k|k-1} \\
\widehat{\boldsymbol{\varphi}}_{k|k} &= \widehat{\boldsymbol{\varphi}}_{k|k-1} + \mathcal{M}_\infty (\mathbf{s}_\alpha - \widehat{\mathbf{s}}_{k|k-1}) \\
\widehat{\boldsymbol{\varphi}}_{k+1|k} &= \mathcal{A}_\varphi \widehat{\boldsymbol{\varphi}}_{k|k} \\
\widehat{\boldsymbol{\phi}}_{k+1|k}^{(\beta)} &= \mathbf{H}_\beta \widehat{\boldsymbol{\varphi}}_{k+1|k}.
\end{aligned} \tag{3.123}$$

The prediction model,  $\mathcal{A}_\varphi$ , is a block-diagonal matrix composed of SA  $\mathcal{A}_\delta^*(l)$  which predict the phase in each layer at time-step  $k + 1$ ; it is computed using the same principle as the SA LQG prediction model in Eq. 3.118, but over the larger meta-pupil of each layer,

$$\mathcal{A}_\varphi = \Sigma_{\varphi+\delta, \varphi} \Sigma_\varphi \tag{3.124}$$

$$= \begin{bmatrix} \Sigma_{\varphi_{v_1 T_s}, \varphi_0^1} \Sigma_{\varphi_0^1, \varphi_0^1}^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \Sigma_{\varphi_{v_n T_s}, \varphi_0^n} \Sigma_{\varphi_0^n, \varphi_0^n}^{-1} \end{bmatrix} \tag{3.125}$$

where  $\varphi_0^i$  is the meta pupil phase in layer  $i$  directly above the pupil, and  $\varphi_{v_i T_s}$  is the meta-pupil phase in layer  $i$  shifted by the Taylor frozen flow approximation by the distance equal to the layer wind velocity times the sample period.

## Chapter 4

# Results of Tomographic Wavefront Reconstruction Simulations of Raven

System modelling of Raven was carried out in two stages. Stage one involved the study of a broad swath of parameter space using two independent simulation platforms. The parameter space was explored in order to establish and/or verify design parameters, as well as make the initial determination that Raven can realistically achieve the proposed performance requirements and deliver useful MOAO-corrected images to the Subaru IRCS spectrograph. Stage two focussed specifically on implementing the tomographic reconstructor algorithms developed in Chapter 3. The intentions of stage two were three-fold: fully understand the expected performance and system limits in the baseline case; gain a platform-specific estimate of the computational complexity required both to execute each reconstructor in the AO loop and compute them in the background, and finally, evaluate the relative performance of each reconstructor. The most promising algorithms, in terms of performance increase, were selected and ported to the Raven system in order to test with the telescope simulator.

This chapter begins with a full description of the end-to-end simulation of the Raven system, including a series of validation tests which confirm that the simulation tools are performing as expected. A summary of the stage one parameter space study follows. Finally the results of the stage two reconstructor study are presented in detail. Some extensions based on the specific optical design of Raven are included at the end of the chapter.

## 4.1 End-to-end simulation

The OOMAO (Object Oriented MATLAB Adaptive Optics) modelling library, a MATLAB-based AO simulation tool [89], is a set of MATLAB classes developed for the purpose of facilitating a clear and accessible end-to-end model of an AO system. Objects from the different classes are assembled to perform diffractive numerical modelling of the optical components of an AO system. OOMAO can be seen as an extension of the MATLAB language; overloaded operators are used to propagate the wavefront through the system and to update the status of each object. Asterisms can be defined, using the source class, with any number of guide stars in constellations specified by polar coordinates. The source class has a very important role in the OOMAO library as it is the link between other classes. A source object carries a wavefront, both amplitude and phase, through the different objects representing the atmosphere, the telescope, the wavefront sensor, etcetera. Both NGS and LGS asterisms can be simulated. Currently all guide stars in the same asterism are defined with the same magnitude and wavelength. Science source objects can be defined individually, assigned their own magnitude and wavelength, and placed at any point in the FoR. Using these classes, an end to end open loop model of Raven has been developed with movable science objects in an adjustable asterism.

In all simulation results that follow, a thresholded CoG algorithm was used to measure the spot positions on the WFSs, although more sophisticated centroiding algorithms, such as correlation centroiding, are planned for implementation on Raven in the future.

### 4.1.1 Temporal properties

Tests were performed to ensure the statistics of the model were behaving as expected. This included a study of the temporal decorrelation of Zernike modes in the simulated atmosphere object. The theoretical temporal decorrelation of several modes was computed as described in Sec. 3.3.2 and compared to the temporal covariance computed from the modal decomposition of a long series of simulated phase measurements taken at different sampling rates. Initially, the temporal behaviour of the simulated phase was not sample-rate independent and did not match the theoretical curves (Fig. 4.1).

The phase screens are generated as needed based on sampling time, wind speed, wind direction and pixel scale. It is done using the method described in [80]. It was thought that there might be a fundamental limitation in the method, however

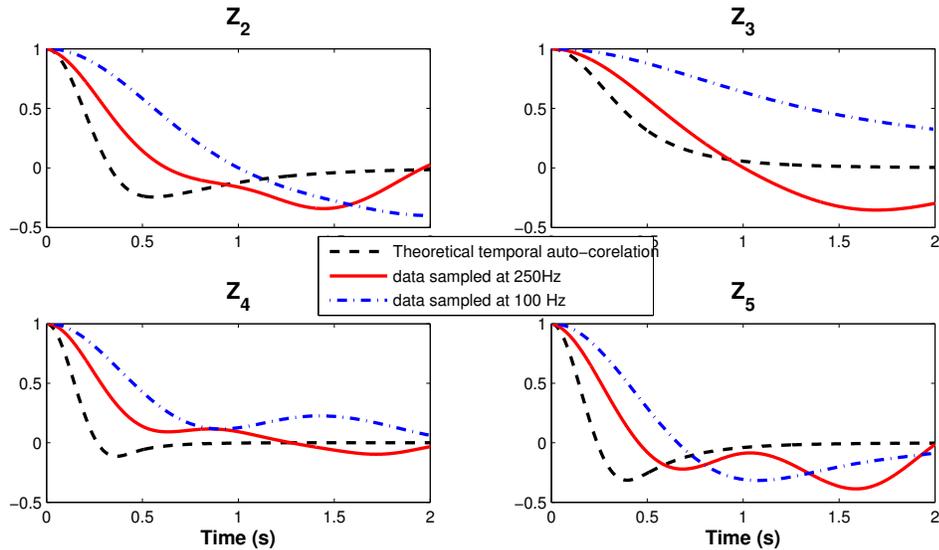


Figure 4.1: Atmosphere simulated phase temporal auto-correlation compared to theory: Does not match.

fixing an error within the implementation handling interpolation of fractional pixel translations of the phase resolved the issue, leading to good agreement between theory and simulated data independent of sample rate for the first 0.5 seconds (Fig. 4.2). Differences after that time window likely arise due to the finite set of measurements used to compute the auto-correlation. The temporal prediction is carried out only over intervals up to 50ms where the simulation is a good match to theory.

### 4.1.2 Spatial properties

The spatial covariance of the simulated phase was computed in Zernike space from a large series of simulated data and compared to the theoretical expectation (see Sec. 3.3.1). The values of the diagonal (spatial auto-correlation of each mode) are compared in Fig. 4.3 and it can be seen that the simulation is producing phase screens with the expected spatial properties.

### 4.1.3 WFS linearity

The open-loop nature of an MOAO system means the WFS is seeing the full atmospheric disturbance rather than a residual error. It is a concern that the larger spot motions will push the SH-WFS out of the linear regime, especially as SNR decreases

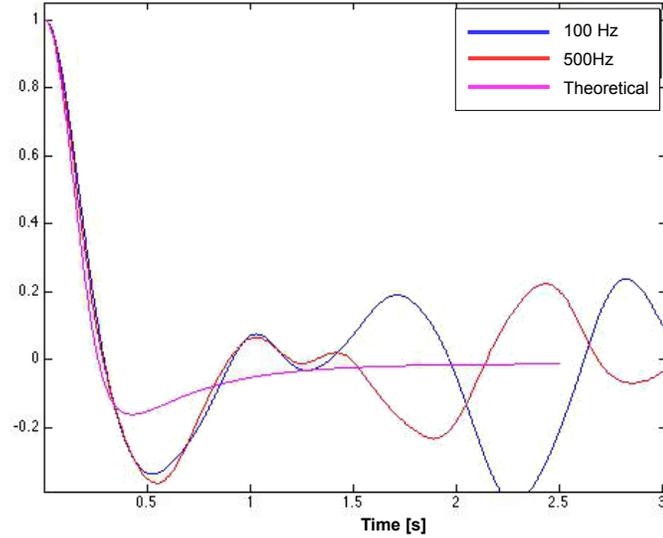


Figure 4.2: Atmosphere simulated phase temporal auto-correlation matches theory for any sample rate after code correction.

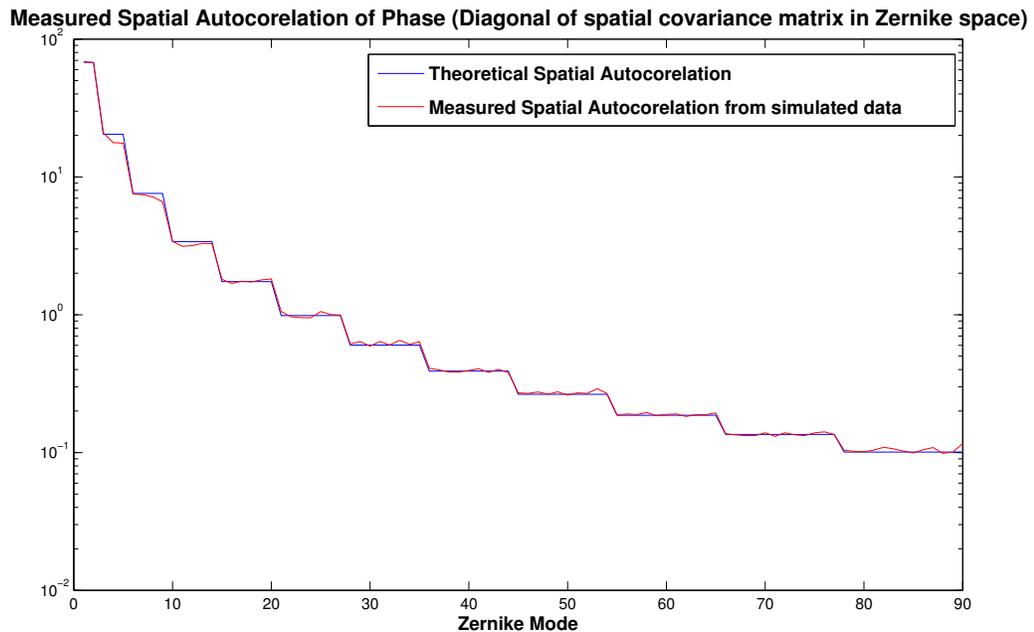


Figure 4.3: Atmosphere simulated phase spatial auto-correlation compared to theory: Good agreement.

for NGSs of increasing magnitude. A simple test, where a tilt of increasing magnitude was introduced to the modelled WFS and the slopes measured for each tilt step, showed decreasing adherence to the linear model with increasing NGS magnitude

(Fig. 4.4a). However, the introduction of a threshold before computing the slopes eliminated the error almost entirely (Fig. 4.4b). As the CoG is almost certainly going to be thresholded on the real system, this is a necessary parameter and is realistic to use.

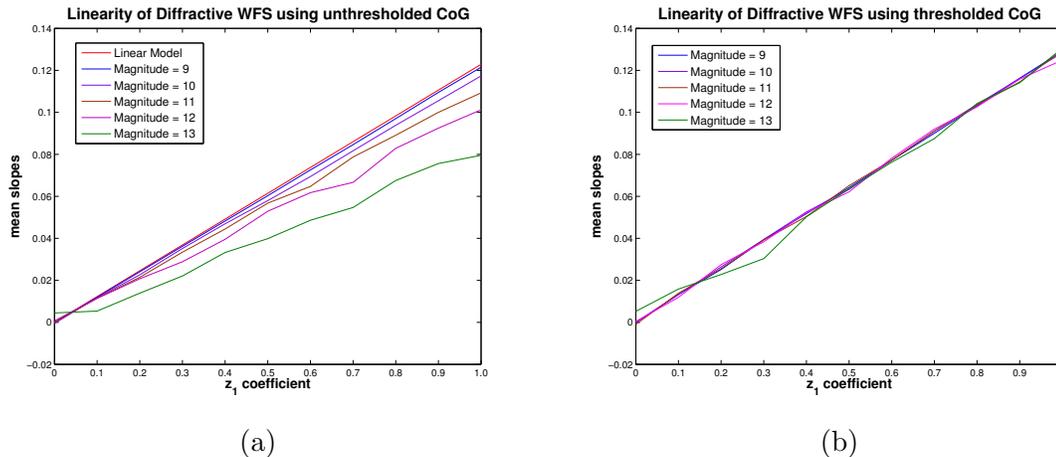


Figure 4.4: Linearity test of the diffractive model of a SH WFS as a function of magnitude without threshold (left) and with threshold (right).

#### 4.1.4 WFS noise

Photon noise and read-out noise are added to the simulated WFS camera frame at read-out. The photon noise has Poisson statistics, scaled by an excess noise factor of  $\sqrt{2}$  to reflect the photon noise characteristics of the EMCCD cameras used in Raven. Read-out noise is kept low by these cameras and is set between 0.2 and 0.5 events. The simulated WFS spot images are shown in Figs. 4.5a and 4.5b for a bright star (magnitude 0) and a dim star (magnitude 13).

#### 4.1.5 Science images

Verification of the science images was carried out by measuring the FWHM of a long exposure of an uncorrected science object on the simulated imager. The theoretical estimate of the FWHM of the long exposure PSF for the baseline atmosphere in J band ( $1.2 \mu\text{m}$ ) can be computed for a finite outer scale using the expression developed in [90] for the ratio,  $L_0/r_0 > 20$ . Starting from the expression of FWHM for infinite

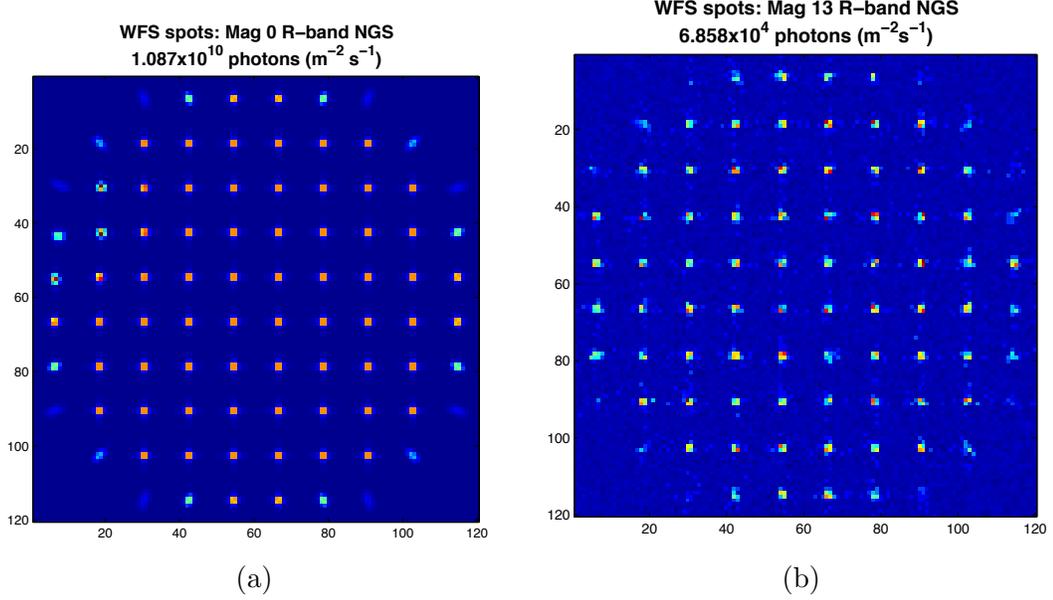


Figure 4.5: Simulated WFS noise properties and photon counts.

outer scale,

$$\epsilon_0 = 0.98 \frac{\lambda}{r_{0,\lambda}}, \quad (4.1)$$

where the Fried parameter given at  $0.5 \mu\text{m}$  can be scaled with wavelength according to the 6/5 power law,

$$r_{0,\lambda} = r_0 \left( \frac{\lambda}{0.5} \right)^{(6/5)}, \quad (4.2)$$

the modified FWHM due to the atmosphere given a von Kármán model is given as,

$$\epsilon_{vK} = \epsilon_0 \sqrt{1 - 2.183(r_0/L_0)^{0.356}}. \quad (4.3)$$

With  $r_0 = 0.156 \text{ cm}$  at  $0.5 \mu\text{m}$  and  $L_0 = 30 \text{ m}$ , this computation yields an expected FWHM of 0.389 arc seconds in J-band. A two dimensional gaussian was fit to a simulated long exposure science image in J-band from which the FWHM in pixels was extracted and converted to arc seconds by multiplying by the applied pixel scale in the simulated imager. The result of this computation was a simulated FWHM of 0.387 arc seconds, confirming good agreement between theory and simulation.

## 4.2 Parameter space study

The main focus of this study was setting the basic system parameters, such as system order, field of regard, and the limiting magnitude. To understand how performance varies with these and other parameters, we defined an asterism with 3 NGS on a ring with a 45 arc-second radius and evaluated the performance at multiple field points within that circle (in some simulations a LGS WFS was included at the field centre). The average performance is defined over points out to 30 arc-seconds from the field centre (excluding the field centre when a LGS was used in the simulation).

The analysis was undertaken using two independent simulation platforms: MAOS (Multi-threaded Adaptive Optics Simulator) and OOMAO. MAOS is an AO simulation tool developed in C by Lianqi Wang and TMT. A more detailed description of the MAOS code and the results of the parameter space study are available in [51] in Appendix A.

The figure-of-merit used to establish the quality of the wavefront correction is the EE, SRs are also computed, as it is a value of interest to the astronomical community, both quantities are described in Sec. 2.1.4.

As an initial validation, a comparison of results between the two simulation tools was made. The predicted Raven performances from the MAOS and OOMAO simulations of the baseline system (Table 4.2) are in excellent agreement. Both simulation tools predicted a mean Strehl ratio of 30% for the points within 30 arc-seconds of the field centre. The two simulations also predicted an identical 43% ensquared energy in this area. Only the mean wavefront error (WFE) differed slightly; MAOS predicted a slightly lower mean WFE of 290 nm while OOMAO predicted 300 nm rms of WFE. The minimum WFE, reached at the NGS radius, is 220 nm rms. This is higher than the 180 nm rms best case WFE obtained with MAOS. It is suspected that either the thresholded CoG used by OOMAO gives slightly poorer performance than the matched filter used in MAOS under these conditions, or the finite number of radial orders (8 in this case) of the static modal spatio-angular reconstructor, used as the preliminary test case, for doing tomography with OOMAO, limits the simulated performance. Overall, however, the excellent agreement between these two independent AO simulation tools provided confidence in our results and signalled that both tools can be used interchangeably in Raven simulations. The subsequent studies of tomographic reconstructions were carried out in OOMAO.

The TMT MASS/DIMM site survey data from site MK 13N [91] (on Mauna Kea)

Table 4.1: Subaru Model Profiles for Raven

h (km)	Fractional Layer Strength		
	25%	50%	75%
$r_0$	19.4 cm	15.6 cm	12.1 cm
0	0.6823	0.5960	0.4971
0.5	0.0611	0.0963	0.1382
1	0.0212	0.0325	0.0577
2	0.0172	0.0372	0.0642
4	0.0757	0.0869	0.0833
8	0.0486	0.0684	0.0895
16	0.0939	0.0826	0.0700

was used to assemble a realistic multi-layer profile. The TMT site survey group created 7 layer profiles by sorting the thousands of individual MASS/DIMM profiles by the calculated  $\sigma_{fit} + \text{lag error}$  (uncorrectable residual WFEs). While this quantity is not equivalent to image quality (IQ), it is an acceptable representation. To assemble representative profiles corresponding to the quartiles of this residual WFE, 10% of the profiles clustered around the quartiles were averaged.

We complemented these MASS/DIMM profile measurements with measures of Subaru IQ taken between 2000 to 2004 [92]. The image-quality derived  $r_0$  values are smaller than the TMT site testing  $r_0$  measurements. One can interpret these differences as being due to the local ground layer at the Subaru telescope, windshake, and dome seeing. Assuming an infinite outer scale, we determined the Fried parameter,  $r_0$ , for each quartile of IQ, generating 3 new profiles with much stronger ground layer components (Table 4.1).

The added dome/ground seeing component in these profiles skews the fraction of the turbulence below 1 km to be greater than 60%. We simulated the performance of Raven using the median Subaru profile defined above. The 3 NGS were placed on a 45 arc-second radius ring and the WFSs have 10 subapertures across the pupil. The performance was sampled across the FoR out to 60 arc-seconds, as shown in Fig. 4.6.

The results for that system using a static modal reconstructor is shown in Fig. 4.7. A system order trade study outlined in [51] concluded that little improvement in performance can be gained (with Raven) for higher order systems. A study of asterism diameter also showed that there is limited improvement on an 8m telescope for guide

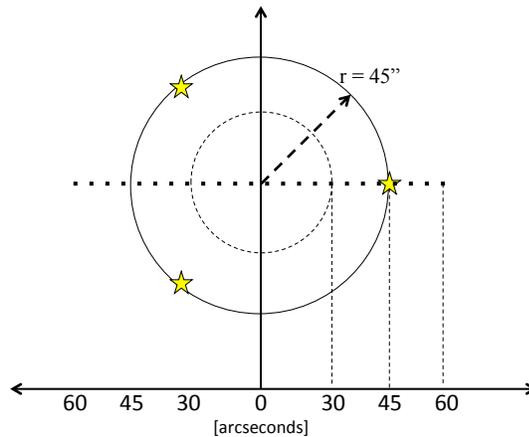


Figure 4.6: Baseline test configuration: A regular asterism with a radius of  $45''$ , performance was sampled at regular intervals from on-axis out to  $60''$  in opposite directions. Average performance is computed from samples out to  $30''$  regardless of asterism radius.

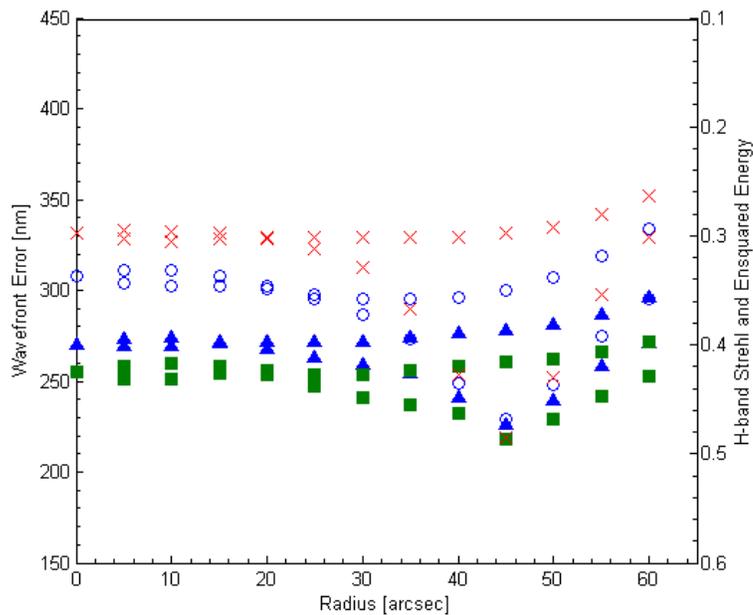


Figure 4.7: Performance for baseline system. Blue triangles are rms WFE from full error signal, green squares are T/T removed rms WFE (left axis). Open circles are EE and red X's are Strehl ratio (right axis).

star separations greater than 2 arc-minutes diameter, as the footprints of the NGSs in the metapupils separate at relatively low altitudes, leaving much of the atmospheric

turbulence unsensed and uncorrectable, while reducing the asterism diameter clearly reduces the field over which targets can be selected. Correction can still be carried out when a target is outside the asterism, but the narrower the asterism, the faster performance falls off. The performance as a function of Science target position within asterisms of increasing diameter is shown in Fig. 4.8. Here it can be seen that a 2 arc-min asterism meets or exceeds the 30% EE requirement over the entire sample range, using a larger asterism means there will be a reduced fraction of the sky enclosed by the asterism over-which the minimum performance can be met.

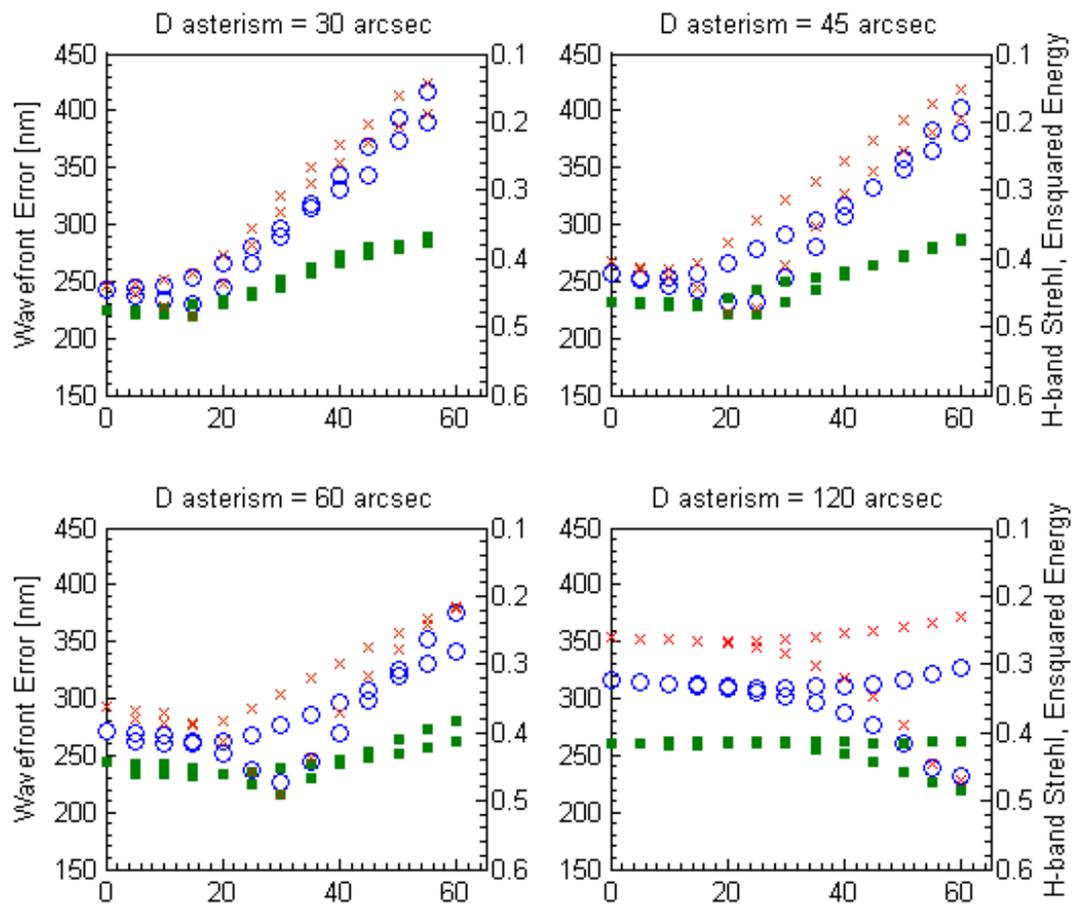


Figure 4.8: Performance for asterisms of varying diameter: Green squares are WFE (left axis). Open circles are EE and red X's are Strehl ratio (right axis).

## 4.3 Modal reconstructors results and analysis

A series of numerical simulations will show the improvement in system performance, in terms of EE and SR, starting from the static and predictive modal reconstructors in Zernike space developed in Secs. 3.3.1 and 3.3.2, leading up to the fully optimal LQG controller presented in Sec. 3.3.4. A concise summary of the results and formulations of the static and predictive modal reconstructors is available in [58].

A systematic survey of the performance trade-off as a function of the temporal lag and SNR parameter space was undertaken in order to establish the peak performance achievable for each type of reconstructor for NGS magnitudes increasing from  $R = 13$  (bright in this context) to  $R = 17$ . This is done with the goal of confirming the preliminary assessment given in Sec. 3.3.3 that carrying out temporal prediction using an AR2 model can lead to an effective increase in the system's limiting magnitude by half a magnitude and the SA model can provide an increase of one magnitude. The study will also provide an estimate of how much the LQG algorithm can be expected to improve the limiting magnitude.

Results of Monte Carlo numerical simulations show the system performance using EE and SR as the benchmark figures of merit. Simulation parameters were selected to reflect the laboratory observing conditions of Raven with the telescope and atmosphere simulator. This includes an asterism of 3 NGSs within a 2.5 arcmin FoR. In this case, an asterism with a radius of 0.5 arcmin was selected in order to reduce tomographic errors and highlight the temporal aspects of system performance. The full set of simulation parameters are listed in table 4.2.

### 4.3.1 Static reconstructors: Verifying the equivalence of explicit and SA formulations

The primary objective in developing the static reconstructors was to demonstrate the equivalence of the explicit layer estimation and the pupil-plane SA estimation. The explicit estimate of the phase in the atmospheric layers provides a wealth of information that is useful for the more complex prediction algorithms; however, the real time computational load scales with the number of layers in the atmospheric model. By contrast, the real-time complexity of the SA reconstructor remains identical regardless of the number of atmospheric layers, only increasing the computational load of the background task required to generate the reconstructor itself. There is, therefore,

Table 4.2: Raven Baseline Configuration Parameters used in all modal simulation cases presented below.

<b>Telescope</b>	
D	8 m
<b>Atmosphere</b>	
$r_0$	15.6 cm
$L_0$	30 m
zenith angle	0 deg
Fractional $r_0$	[0.596; 0.224; 0.180]
Altitudes	[0; 5;10] km
wind speeds	[7.5; 12.5; 15] m/s
wind direction	[0; 0; 0] deg
<b>Wavefront Sensor</b>	
RON	$0.2 e^-$
$N_{NGS}$	3
NGS radii	30 arcsec
Order	$10 \times 10$
$\theta_{pix}$	0.4 arcsec
$N_{pix}$	15
$f_{sample}$	30-200 Hz
$\lambda_{WFS}$	$0.64 \mu\text{m}$
Centroiding algorithm	thresholded Centre-of-Gravity
<b>DM</b>	
Order	$11 \times 11$
stroke	infinite
influence	cubic
<b>AO loop</b>	
pure delay	3ms
controlled modes	$N_z=55$
reconstructor	$N_p= 406$ modes
predictor	90 modes
<b>Evaluation</b>	
$\lambda_{evl}$	$1.65 \mu\text{m}$ (H-band)

a trade-off which exists between model accuracy, computational complexity and the amount of information generated by the algorithm.

The simulation results confirm the theoretical computation made in Sec. 3.3.1 which indicated that the performance of the explicit layered reconstructor would approach that of the SA reconstructor as the number of radial orders used to generated

the component covariance matrices was increased. The measured EE as a function of radial order for the baseline test parameters (with an asterism diameter of 1.5 arc-minutes) is shown in Fig. 4.9. As predicted, the error becomes very small at 27 radial orders and becomes negligible at 36 radial orders.

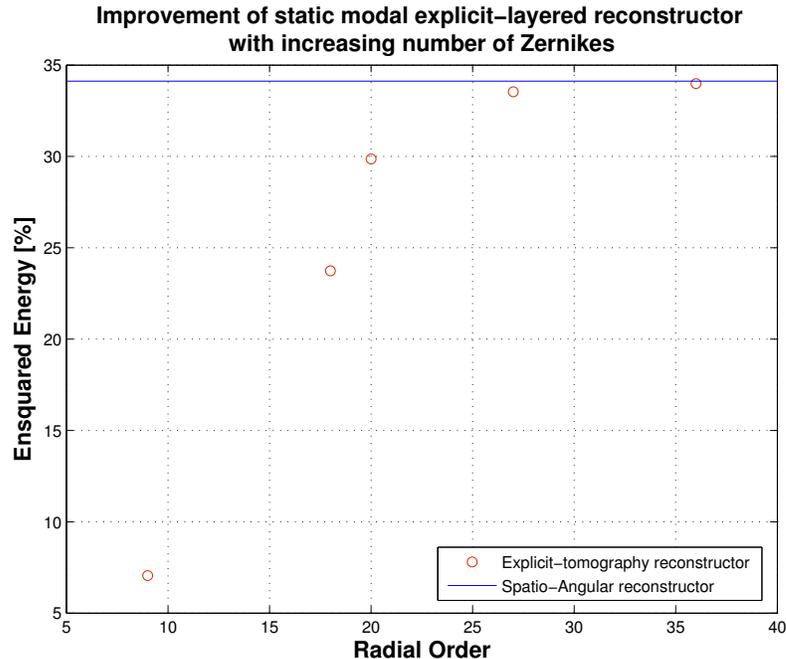


Figure 4.9: As the number of Zernike modes used to compute the explicit layer reconstructor increases, system performance approaches that of the spatio-angular reconstructor which is computed using an analytical expression and effectively accounts for an infinite number of modes.

### 4.3.2 Comparison of first and second order predictive models

Based on the analysis carried out in Sec. 3.3.2, the AR2 and SA prediction models were implemented in simulation and their performance compared to the static reconstructor at various frame rates and NGS magnitudes. The results of the numerical simulations are shown in Fig 4.10 and summarized in table 4.3. Following sec. 3.2.4 turbulence is expanded on 27 radial orders (i.e. 406 modes) from which 9 radial orders (55 modes) are controlled. Because the parameter identification is carried out mode-by-mode, despite parallelization and the speed up achieved by the minimization described in Sec. 3.3.2, the computational load is high. As such, the predictive

models are computed for 90 out of 406 modes per layer with no significant effect on the overall performance.

The data show the peak system performance in SR and EE, on the Subaru IRCS 140 milliarcsecond slit, for a 30 arc-second radius asterism of 3 guide stars of the same magnitude, and the corresponding sampling rate at which the peak occurs for each of three algorithms: Static SA, 2-step predictive using an AR2 model, and 1-step predictive using the SA predictive model. For interest and comparison purpose, the AR1 model was also tested. Lags corresponding to frame-rates of 30 to 200 Hz were tested. The simulation incorporates the fixed system lag of 3ms allotted for camera readout, data processing and issuing of DM commands. In order to model the delayed application of the DM commands in the middle of an exposure, the system is attributed a fixed sample rate of 1kHz. The resulting output phase is summed over the first 3 milliseconds (three samples) before the new DM command is applied; subsequently, the output phase for the remaining total exposure time is added to these first frames to make a total exposure of the desired length. Each simulation run collected 2000 exposures before computing the SR and EE.

The results confirm that a reduction in frame-rate, combined with SA temporal prediction – as noted in [60] – will allow the system to achieve a level of performance for a given GS magnitude which is equal to the performance with a static reconstructor using GSs one magnitude brighter, and the performance with an AR2 predictive reconstructor using GSs half a magnitude brighter. The results confirm the computations shown in Fig. 3.7 that estimate no reduction in temporal error with the use of an AR1 prediction model.

The simulated science image obtained using the static reconstructor is shown next to that obtained using the SA prediction model in Fig. 4.11 when using NGSs of magnitude 17 and OL-WFSs operating at 25Hz.

### 4.3.3 LQG results and analysis

Based on the results obtained with the different prediction models in Sec. 4.3.2, the LQG controllers introduced in Sec. 3.3.4 were implemented in Zernike space. First, the AR2 prediction model was used in a two-state controller and an improvement in SR indicates an increase in limiting magnitude by up to 1.5 magnitudes. The increase in EE gives an equivalent performance at magnitude 17 as that obtained with a static reconstructor at magnitude 14.

Table 4.3: Raven End-to-End simulation results. The optimal performance ( % en-squared energy) for each GS magnitude is shown for three reconstructors: the static MMSE, Spatio-Angular prediction and the AR2 prediction model. AR1 prediction is included for comparison purposes

GS mags (R)	static MMSE				AR2 prediction			
	EE	lag [ms]	Strehl	lag [ms]	EE	lag [ms]	Strehl	lag [ms]
14.0	35.70	7	30.08	7	35.83	9	30.71	10
14.5	35.39	7	28.70	9	35.65	10	29.89	11
15.0	35.32	9	26.69	10	35.38	10	28.38	12
15.5	34.48	10	24.39	12	35.27	12	26.85	14
16.0	33.61	12	21.58	14	34.12	14	24.23	17
16.5	32.24	14	18.05	17	33.59	18	21.60	21
17.0	30.39	18	14.39	27	32.16	21	17.91	28
	SA prediction				AR1 prediction			
	EE	lag [ms]	Strehl	lag [ms]	EE	lag [ms]	Strehl	lag [ms]
14.0	36.70	9	32.26	10	35.67	8	30.09	8
14.5	36.41	9	30.98	10	35.59	9	28.73	9
15.0	36.00	9	29.74	12	35.11	9	26.71	12
15.5	35.49	12	28.47	17	34.38	10	24.32	13
16.0	35.34	23	26.95	24	33.30	12	21.29	14
16.5	35.10	27	24.81	27	31.98	14	17.97	18
17.0	34.24	33	21.41	33	30.27	19	14.30	28

The success of the stand-alone SA prediction model also translated into improved performance within the LQG algorithm, making an incremental gain over the LQG with the AR2 model. Overall, it increased both Strehl and EE values well above the performance of the static reconstructor. Recall that both the modal SA and AR2 LQG carry out an estimate of the predicted phase in each layer of the atmosphere model before propagating the estimated phase into the pupil plane in the direction of the science objects, but the benefit of the SA prediction model is that it only requires the state vector to contain the phase vector from a single time step, reducing the size of the state vector by a factor of 2. The simulation results for the LQG implemented with each prediction model are summarized in Table. 4.4 and Fig. 4.12, which again show the best performance for each algorithm at increasing GS magnitudes and the sample time at which that performance was obtained.

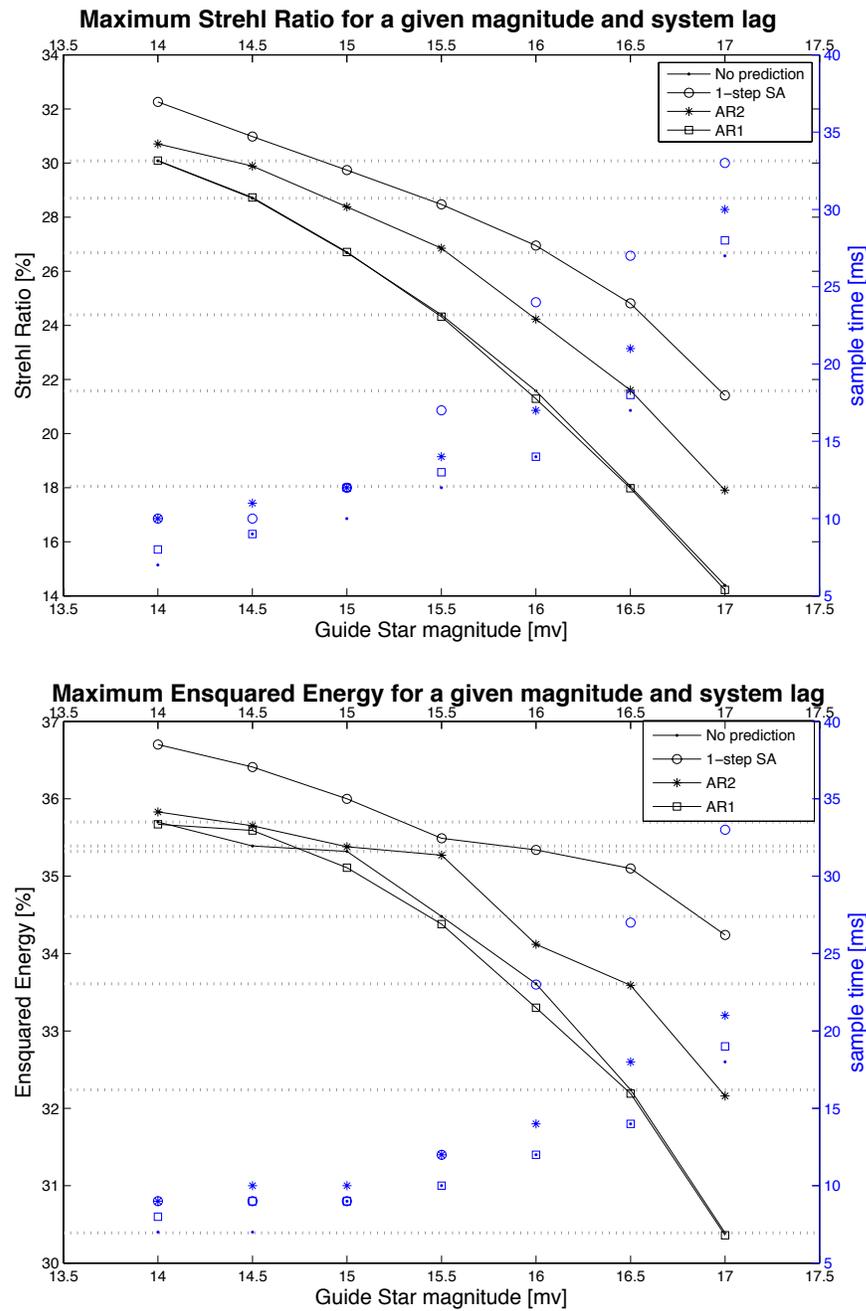


Figure 4.10: Simulation results showing peak values of Strehl ratios (top) and ensquared energy (bottom) and the WFS framerate at-which they occur as a function of tomographic algorithm for each magnitude using stand alone prediction models AR1, AR2 and 1-step SA.

#### 4.3.4 Meeting Raven science requirements

The numerical simulation results presented in the sections above were carried out with tight asterisms ( $r = 0.5'$ ) in order to reduce the tomographic error and highlight

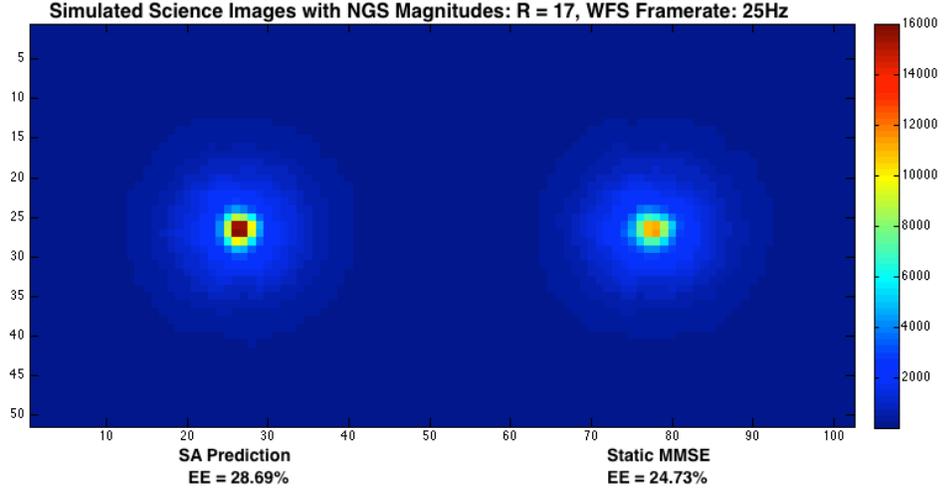


Figure 4.11: The simulated science images for the SA predictive model (left) and the static reconstructor (right) in modal space.

Table 4.4: Raven End-to-End simulation results. The optimal performance ( % en-squared energy) for each GS magnitude is shown for the LQG reconstructor using the Spatio-Angular prediction model and the AR2 prediction model.

GS mags (R)	LQG with AR2				LQG with SA			
	EE	lag [ms]	Strehl	lag [ms]	EE	lag [ms]	Strehl	lag [ms]
14.0	43.05	10	36.90	10	45.36	10	41.42	15
14.5	42.56	11	35.28	15	44.84	16	40.63	22
15.0	42.30	11	33.59	14	44.58	15	40.01	23
15.5	41.22	12	30.79	15	43.98	17	37.37	22
16.0	39.57	15	27.36	19	43.58	25	36.88	27
16.5	37.84	16	23.63	24	42.99	30	34.01	33
17.0	35.26	20	19.41	29	41.89	36	30.12	35

the effects of reduced temporal errors. The original science requirement for Raven states that an EE of 30% should be achieved in H-band in the baseline case, that is for an  $r_0$  of 0.156m and an asterism diameter of 1.5 arc-minutes. The initial trade study in [51] concluded that this could be met for NGSs no dimmer than magnitude 14.5 using a static reconstructor and thresholded CoG centroiding algorithm at a reduced framerate of 180Hz. This was confirmed by the current model and a composite plot (Fig. 4.13) of system performance with increasing lag at each magnitude for each type of reconstructor shows that simple prediction (AR2 model) will bring performance with magnitude 15 NGSs up to the limit, and LQG with SA prediction will bring

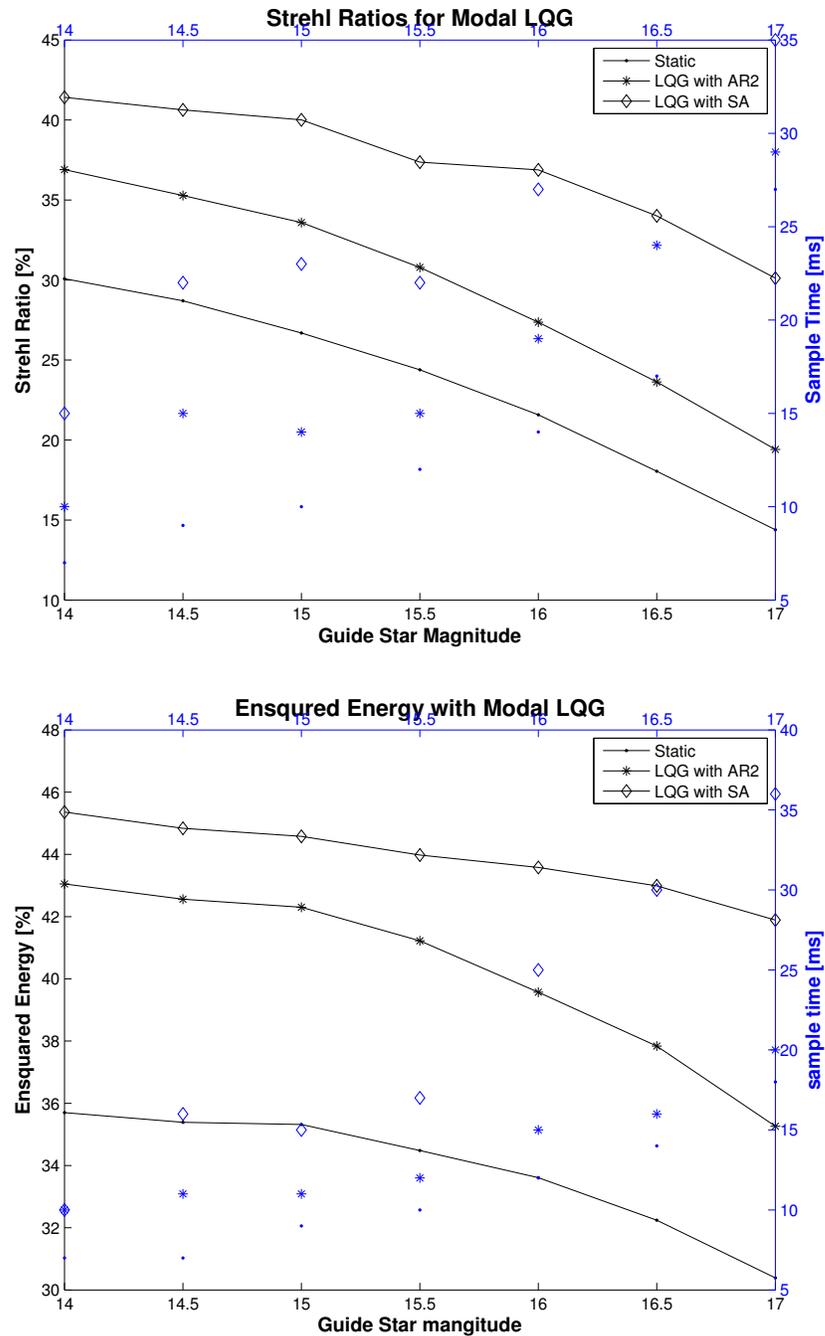


Figure 4.12: Static modal reconstructor is replotted against the modal LQG algorithm simulation results using the AR2 prediction model and the SA prediction model.

performance with magnitude 17 NGSs well over the limit, confirming that the results obtained with low tomographic error hold under less favourable conditions.

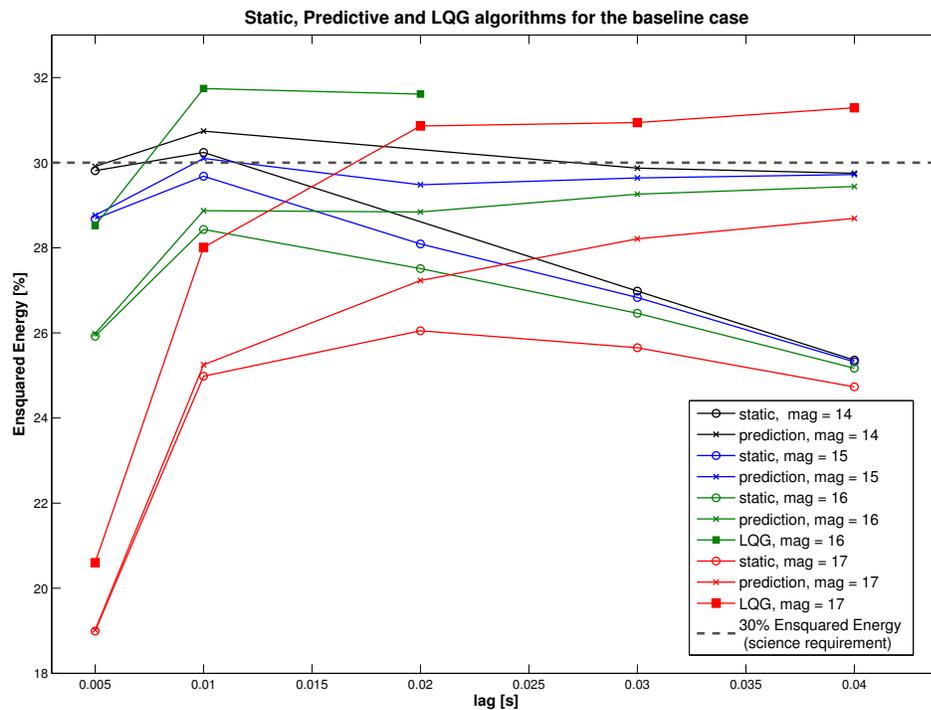


Figure 4.13: The use of more complex predictive algorithms will allow Raven to meet scientific requirements using dimmer NGSs therefore increasing sky coverage.

## 4.4 Zonal basis results and analysis

The Zernike basis is an intuitive and well established space for WF reconstruction, however the need to take into account a large number of modes above those being corrected to avoid aliasing, and other long computational steps in the background process make the modal reconstructors largely impractical to implement on Raven. It is certainly possible to streamline the background tasks through efficient programming, but that was not the primary focus of this project. The zonal reconstructors described in Sec. 3.4 are overall much quicker to compute. As was shown in Sec. 3.4.4, a very minimal implementation of the zonal LQG was established by taking advantage of the specific architecture of MOAO (this could also be developed in modal space but the high background computational complexity would remain). The static,

predictive and LQG algorithms in zonal space were implemented in simulation and the results are shown here.

#### 4.4.1 Static reconstructor

The static zonal SA reconstructor has the benefit of being much faster to compute than its counterpart in the modal basis. It was also found in simulation that the performance of the zonal representation slightly exceeds that of the modal representation, even when accounting for the spatial aliasing in the computation and reconstructing the maximum number of radial orders that can be estimated in the system (Fig. 4.14).

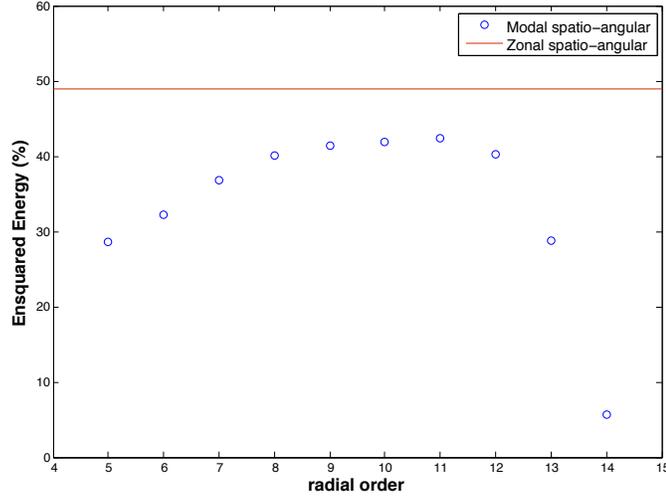


Figure 4.14: Static reconstructors in zonal and modal basis. A discrepancy in performance can be traced to modal aliasing.

It was later established that the cause of this mis-match was the non-analytical measurement model,  $\mathbf{\Gamma}$  used in simulation to convert slopes to Zernike coefficients. A modal DM calibration step was used, leading to spatial aliasing, where an analytical expression from the derivatives of Zernike polynomials would have been more accurate. When this analytical expression is implemented, the overall performance of the two reconstructors matches very well.

A numerical comparison of the Explicit,

$$\hat{\mathbf{s}}_{\beta} = \mathbf{\Gamma} \mathbf{H}_{\beta} \mathbf{\Sigma}_{\varphi} \mathbf{H}_{\alpha}^{\top} \mathbf{\Gamma}^{\top} (\mathbf{\Gamma} \mathbf{H}_{\alpha} \mathbf{\Sigma}_{\varphi} \mathbf{H}_{\alpha}^{\top} \mathbf{\Gamma}^{\top} + \mathbf{\Sigma}_{\eta})^{-1} \mathbf{s}_{\alpha} \quad (4.4)$$

and SA versions of the static slopes estimates,

$$\hat{\mathbf{s}}_{\beta} = \mathbf{\Gamma}\mathbf{\Sigma}_{(\beta,\alpha)}\mathbf{\Gamma}^{\top}(\mathbf{\Gamma}\mathbf{\Sigma}_{(\alpha,\alpha)}\mathbf{\Gamma}^{\top} + \mathbf{\Sigma}_{\eta})^{-1}\mathbf{s}_{\alpha} \quad (4.5)$$

show that the maximum difference between the matrices  $\mathbf{H}_{\beta}\mathbf{\Sigma}_{\varphi}\mathbf{H}_{\alpha}^{\top}$  and  $\mathbf{\Sigma}_{(\beta,\alpha)}$  and, likewise the difference between  $\mathbf{H}_{\alpha}\mathbf{\Sigma}_{\varphi}\mathbf{H}_{\alpha}^{\top}$  and  $\mathbf{\Sigma}_{(\alpha,\alpha)}$  is 0.5% in both cases. This can be attributed to the interpolation between phase points taking place in the layers as well as the computation of the covariance matrices,  $\mathbf{\Sigma}_{\varphi}$  and  $\mathbf{\Sigma}_{(\alpha,\alpha)}$  using different amounts of information - in the former case, the number of phase points used in the computation grows with the height of the layer and in the latter case the computation is always carried out on phase points in the pupil. In the Raven end-to-end simulation, these small numerical differences translate into a difference in performance results of  $\Delta SR = 0.32\%$  and  $\Delta EE = 0.36\%$ .

#### 4.4.2 Predictive reconstructor

After an extensive examination of layer-based prediction in modal space, this section will focus strictly on pupil-plane SA prediction in the zonal basis. This is easily justified by a comparison of the component matrices of the predictive explicit reconstructor,

$$\hat{\mathbf{s}}(k+1)_{\beta} = \mathbf{\Gamma}\mathbf{H}_{\beta}\mathbf{\Sigma}_{(\varphi+\delta)}\mathbf{H}_{\alpha}^{\top}\mathbf{\Gamma}^{\top}(\mathbf{\Gamma}\mathbf{H}_{\alpha}\mathbf{\Sigma}_{\varphi}\mathbf{H}_{\alpha}^{\top}\mathbf{\Gamma}^{\top} + \mathbf{\Sigma}_{\eta})^{-1}\mathbf{s}(k)_{\alpha}, \quad (4.6)$$

and the predictive SA reconstructor,

$$\hat{\mathbf{s}}(k+1)_{\beta} = \mathbf{\Gamma}\mathbf{\Sigma}_{(\beta+\delta,\alpha)}\mathbf{\Gamma}^{\top}(\mathbf{\Gamma}\mathbf{\Sigma}_{(\alpha,\alpha)}\mathbf{\Gamma}^{\top} + \mathbf{\Sigma}_{\eta})^{-1}\mathbf{s}(k)_{\alpha}. \quad (4.7)$$

The same amount of numerical difference is seen between  $\mathbf{H}_{\beta}\mathbf{\Sigma}_{(\varphi+\delta)}\mathbf{H}_{\alpha}^{\top}$  and  $\mathbf{\Sigma}_{(\beta+\delta,\alpha)}$  as in the static case, and  $\Delta SR = 0.38\%$  and  $\Delta EE = 0.35\%$  in simulation.

#### First order spatio-angular MMSE prediction in phase space

Only the first order SA predictive model was implemented in zonal space. The results of the simulations comparing the peak performance for each magnitude to the static reconstructor are shown in Fig. 4.15 and summarized in Table 4.5. An interesting trend, also present in the modal case, is the difference between the WFS integration times at which the best EE and the best Strehl ratio are achieved. It is therefore important to understand the figure of merit for a given system in order to optimize

performance.

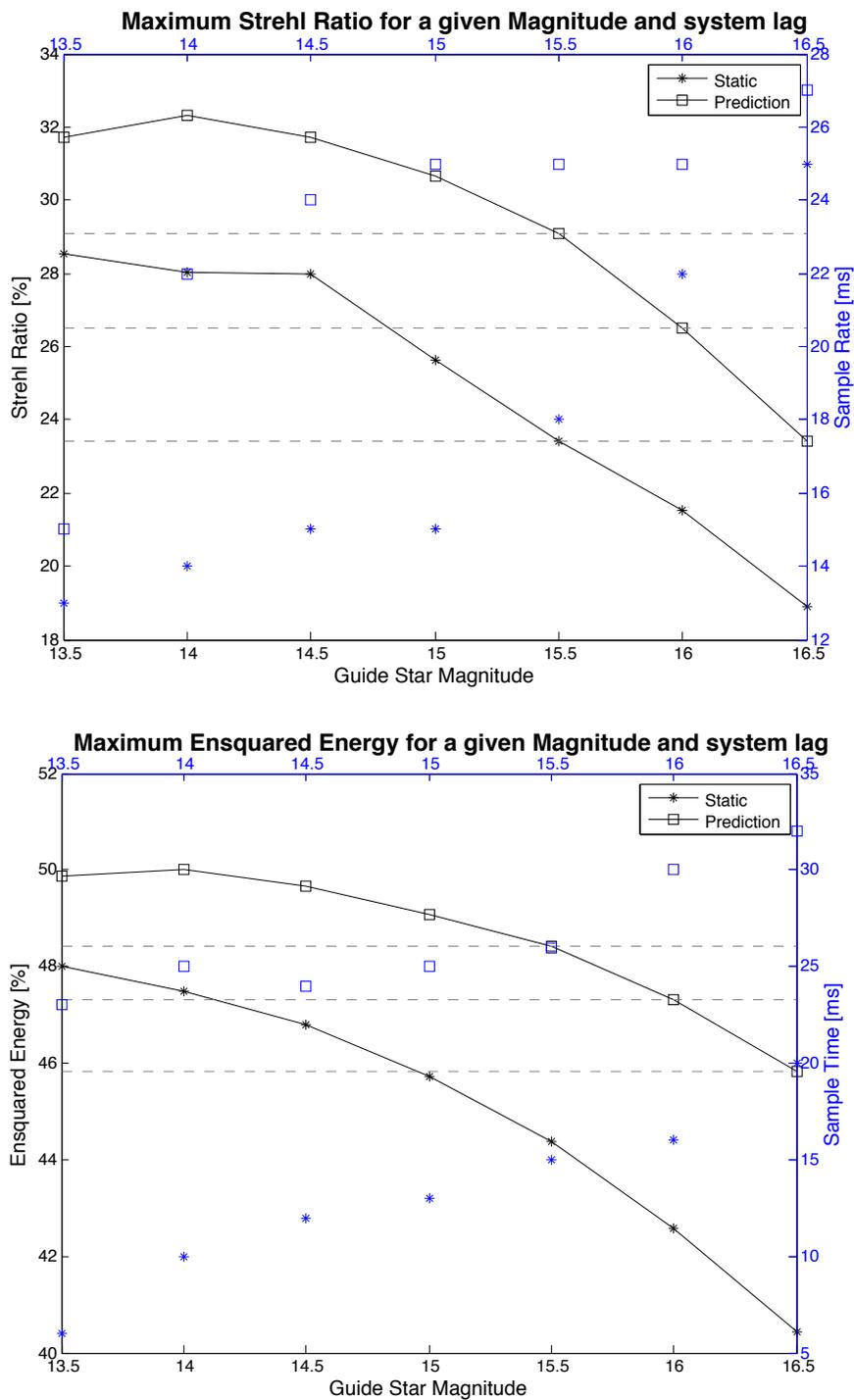


Figure 4.15: Strehl ratios and ensquared energy as a function of algorithm and magnitude using the static zonal reconstructor vs the predictive zonal SA reconstructor.

Table 4.5: Raven End-to-End simulation results. The optimal performance ( % en-squared energy) for each GS magnitude is shown for the zonal static SA, and compared to SA Prediction.

GS mags (R)	static SA				SA Prediction			
	EE	lag	Strehl	lag [ms]	EE	lag [ms]	Strehl	lag [ms]
13.5	47.99	6	28.55	13	49.88	23	31.70	15
14.0	47.48	10	28.05	14	50.01	25	32.32	22
14.5	46.79	12	27.98	15	49.65	24	31.72	24
15.0	45.73	13	25.64	15	49.07	25	30.66	25
15.5	44.36	15	23.41	18	48.41	26	29.11	25
16.0	42.57	16	21.52	22	47.31	30	26.50	25
16.5	40.43	20	18.90	25	45.81	32	23.41	27

### 4.4.3 LQG

#### Zonal LQG in phase space

The results of the numerical simulation of the SA LQG controller as a function of GS magnitudes and WFS frame rates are added to the static and predictive results in Fig. 4.16 and summarized in Tab. 4.6. In the stand-alone prediction model presented above, the spatial and temporal estimates are combined into a single step (Eq. 4.7); in the LQG, the two estimates are necessarily divided into two steps due to the LQG recursion. The two approaches were shown to be mathematically equivalent in Sec. 3.4.2. To verify the equivalence of the one and two-step spatio-temporal estimation, the two methods were compared numerically, with the two-step reconstructor,

$$\hat{\mathbf{s}}_{\beta} = \mathbf{\Gamma} \mathbf{\Sigma}_{\beta+\delta, \beta} \mathbf{\Sigma}_{\beta, \beta}^{-1} \mathbf{\Sigma}_{(\beta, \alpha)} \mathbf{\Gamma}^{\top} (\mathbf{\Gamma} \mathbf{\Sigma}_{(\alpha, \alpha)} \mathbf{\Gamma}^{\top} + \mathbf{\Sigma}_{\eta})^{-1} \mathbf{s}_{\alpha} \quad (4.8)$$

The maximum numerical difference between  $\mathbf{\Sigma}_{\beta+\delta, \beta} \mathbf{\Sigma}_{\beta, \beta}^{-1} \mathbf{\Sigma}_{(\beta, \alpha)}$  and  $\mathbf{\Sigma}_{(\beta+\delta, \alpha)}$  is 0.1%. The effect on performance in the Raven simulation was also minimal with  $\Delta SR = 0.29\%$  and  $\Delta EE = 0.33\%$ .

The plots show that the LQG algorithm can lead to an increase in limiting magnitude of approximately two magnitudes in SR and more than three magnitudes in EE. Another interesting trend is that the best framerate for the LQG and predictive algorithms change very slowly compared to the static reconstructor which changes quite a lot between magnitudes.

Table 4.6: Raven End-to-End simulation results. The optimal performance ( % en-squared energy) for each GS magnitude is shown for the zonal static SA, and compared to SA LQG.

GS mags (R)	static SA				SA LQG			
	EE	lag [ms]	Strehl	lag [ms]	EE	lag [ms]		
13.5	47.99	6	28.55	13	52.27	20	33.54	15
14.0	47.48	10	28.05	14	52.66	22	34.60	22
14.5	46.79	12	27.98	15	52.48	24	34.02	23
15.0	45.73	13	25.64	15	52.06	25	33.12	25
15.5	44.36	15	23.41	18	51.40	25	31.53	25
16.0	42.57	16	21.52	22	50.27	26	29.25	27
16.5	40.43	20	18.90	25	48.98	35	26.72	35
17.0	38.28	25	15.57	25	47.48	45	24.95	45

## 4.5 Robustness

A large set of parameters exist which can be studied, and exploring the entire space in a meaningful way would take a prohibitive amount of time using the simulation tools selected here. Instead, an examination is made of the behaviour of some of the reconstructors under select error conditions that are likely to arise during lab testing and observation. The primary sources of error affecting tomography all stem from using a model of the atmosphere that does not accurately reflect reality. Some of the main culprits will be: inaccurate fractional  $r_0$  profile, mis-identification of layer altitudes and, having particular ramifications for prediction, an incorrect or unknown wind profile. Select cases of these three conditions have been examined in simulation, in some cases using a simplified one-layer system, and in the lab.

### 4.5.1 Robustness to wind profile error

Figure 4.17 shows the analytical temporal error functions for a model mismatch in terms of wind-speed and wind direction. These results indicate that, so long as the wind speed estimate is not negative, the SA predictor can tolerate an error of up to a factor of two before performance degrades to that of the no-prediction case. The results also show that the SA predictor is quite robust wrt wind direction; missing the estimate by  $+/- 50^\circ$  still does better than the no-prediction case.

In consideration of the discussion in Sec. 3.4.3 regarding the estimate of the covariance of the new phase entering the pupil, the bilinear spline interpolation model

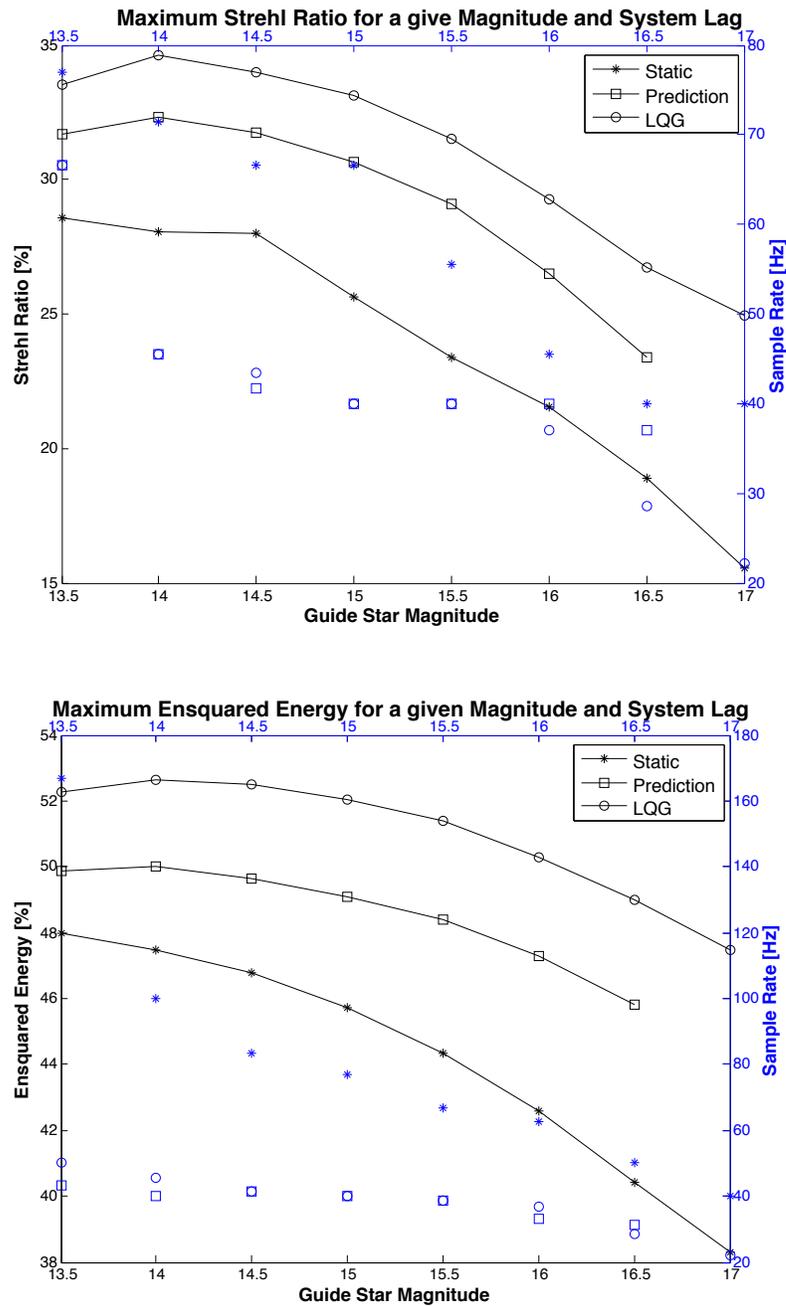


Figure 4.16: Strehl ratios and ensquared energy as a function of algorithm and magnitude for all three static and dynamic zonal reconstructors.

$\mathcal{A}$ , on-which the SA predictor is based, has been built with two versions: one whose new points entering the aperture are interpolated with no further information from within the pupil, thus assuming zeroed wave-front outside the pupil and a model

where for only those new points rows in  $\mathcal{A}$  are kept, the remainder being the two-dimensional bilinear spline weights. Results show that if one does not estimate the new points from a correlation function, the loss in performance is quite drastic with the bilinear spline interpolator achieving best performance for a wind-speed model that is 1/2 the true wind-speed. When the new points' values are computed keeping the corresponding columns of  $\mathcal{A}$  then the performance enhancement is dramatic with some tens of nm rms difference still.

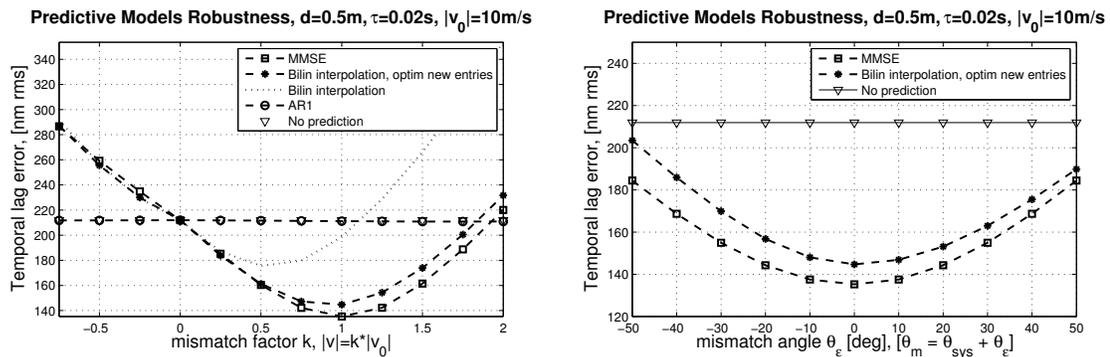


Figure 4.17: (Left) Robustness of zonal SA temporal predictor with respect to wind-speed error compared to other predictive methods. (Right) Robustness with respect to error in wind direction.

#### 4.5.2 Flat $C_n^2$ profile

In the absence of accurate information regarding the nature of the  $C_n^2$  profile, we work under the assumption of no knowledge of the profile weights. In the simulation and in the lab, where the weights on the layers in the input atmosphere are given in Table 4.2 as  $[0.596, 0.224, 0.180]$ , the atmosphere model used to generate the reconstructors is assigned a flat profile with weights,  $[0.35, 0.35, 0.3]$ . An examination of the resulting effect of this assumption on static, predictive and LQG reconstructors in zonal space shows only a small degradation in performance to the predictive and LQG reconstructors and no effect on the static reconstructor. Table 4.7 shows the results for the static and LQG reconstructor where the model uses the known input atmosphere  $C_n^2$  profile compared to the results when the model uses the flat profile.

Table 4.7: Sensitivity of reconstructors to error in the  $C_n^2$  profile estimate.

Method	known $C_n^2$ profile		flat $C_n^2$ profile	
	EE [%]	SR [%]	EE [%]	SR [%]
Static	49.8	31.8	49.7	31.8
Predict	52.6	47.5	51.3	45.6
LQG	54.4	50.1	53.1	48.0

### 4.5.3 Over-modelled atmosphere

The sensitivity of the reconstructors to under-modelling of the atmosphere is a complex question and has been subject to some notable studies [93, 94]. It has been noted in laboratory measurements on Raven, where the input atmosphere is relatively well characterized and known to have 3 layers, that using a fixed, large number of layers (9 in the case of Raven) reduces the sensitivity of the system to model errors even if they are assigned a small fractional  $r_0$  value. It has been shown that this improves performance over a model containing 3 layers with a small amount of error in altitude.

## 4.6 MOAO with CL-WFS correction mode

Taking advantage of the optical design of Raven which has a CL-WFS in each science path, the potential to provide slow CL corrections to the DM has been investigated. Specifically, in the absence of any other static or quasi-static aberration, a simulation was built to establish the CL-WFS framerate at which the injection of un-sensed (by the OL-WFSs) corrections start to degrade the tomographic estimate. Simulating a slow CL-WFS observing a dim (mag 17) but compact science object shows that, in the absence of any drift or quasi-static errors, improved results can be obtained for a framerate between 5 and 50Hz (when the OL framerate is 500Hz) before the performance begins to degrade due to conflict between the OL and CL correction. The control diagram of this AO mode is shown in Fig. 4.18 and the simulation results are given in Table 4.8. The primary intention of this test was to ascertain the effect of low frequency corrections injected on the DM on overall performance to make sure no degradation occurs.

These results indicate that a slow drift or quasi-static error in the science path could be corrected without disrupting the tomographic reconstructor in some science cases with a very slow background update to the DM commands based on long

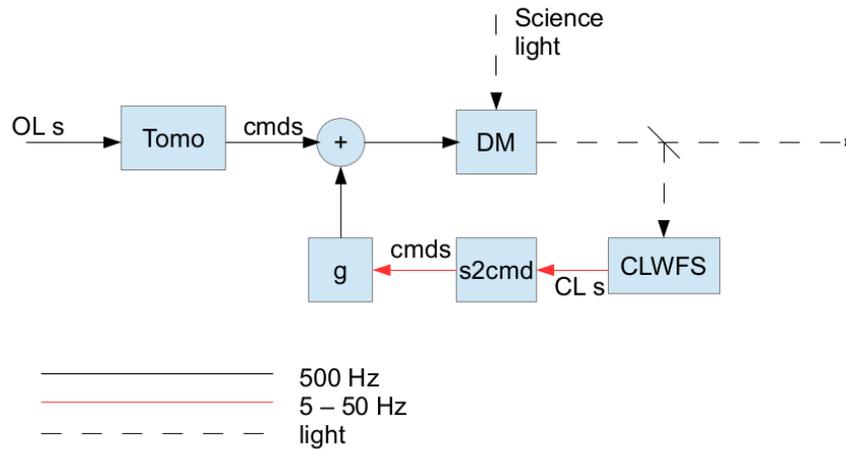


Figure 4.18: The data pipeline when injecting command offsets to the MOAO correction from slow CL-WFS frames from the science targets

Table 4.8: Projected improvement in strehl ratio and ensquared energy when slow CL-WFS measurements of the science target are periodically added as offsets to the real-time (500Hz) tomographic correction.

CLWFS framerate	EE	SR
no CLWFS	43.78	19.11
5 Hz	44.61	21.95
10 Hz	45.79	24.14
25 Hz	45.57	23.50
50Hz	43.92	22.06

exposure CL-WFS frames.

## Chapter 5

# Implementing tomographic reconstruction on Raven

The move from simulation to the real system was carried out in stages, in parallel to the development of the RTC and driving the development of the RPG. The initial testing of the static reconstructors involved snapshot slope measurements with one of the upper layer phase screens in the path in a fixed position. Reconstruction was then carried out off line using the OL-WFS slopes and the quality of the correction was established by comparing the estimated slopes in the science directions to those measured by the CL-WFS. This static testing procedure enabled a procedure to be developed for relating the simulation space to the measurement space; it also provided the first laboratory verification of the equivalence between explicit and SA reconstructors.

This chapter begins with a description of the model identification procedures and system calibration steps directly related to tomography. This is followed by the results obtained in the lab for a cross section of reconstructors both static and dynamic.

### 5.1 System calibration overview

The tomographic reconstructors assume a known system geometry is available. Various steps to ascertain the precise relationship between the model and the system are developed here.

### 5.1.1 Retrieving reconstructor components from the system

Relating the mathematical model space to the measurement space of the system is required to define the geometry of the system. In the case of the tomographic reconstructor, a key element is the spatial location of the GSs and science objects within the FoR relative to each other and relative to some reference coordinate system. In the model, these values are defined by the distance of the objects from the central axis in arc-seconds and the angular distance from an azimuth of zero.

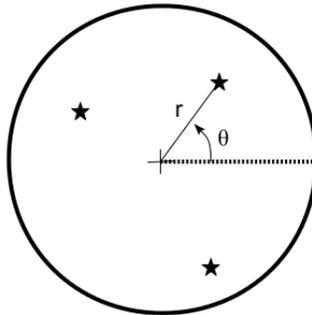


Figure 5.1: Schematic for pinpointing the location of objects in the FoR.

The grid in Fig. 5.2a shows the numbering and location of the calibration unit seeing limited pinhole grid. The distance scaling between the pinholes is 0.532 mm/arc-second and the physical distance is 10 mm. In order to locate the “zero angle” position corresponding to  $\theta = 0$  in Fig. 5.1, a frozen volume of turbulence was setup in the CU - the CDM was set to hold a given shape and the phase screens were inserted and left stationary. A single OL-WFS was moved to pick up the on-axis pinhole which is assumed to have location  $(0,0)$ , and a measurement of the static perturbation was taken with that WFS at that location ( $\mathbf{s}_\alpha$ ). One of the science arms was moved to a pinhole near the central pinhole and a measurement of the static perturbation was taken with the CL-WFS at this location ( $\mathbf{s}_\beta$ ). Holding all other parameters fixed, the reconstructor,  $\mathbf{R}$ , with the best pinhole location parameters, given as  $(\mathbf{r}_i, \boldsymbol{\theta}_i)$ , should minimize the difference,

$$\mathbf{R} = \arg \min_{\mathbf{r}_i, \boldsymbol{\theta}_i} \langle \|\mathbf{s}_\beta - \mathbf{R}(\mathbf{r}_i, \boldsymbol{\theta}_i)\mathbf{s}_\alpha\|^2 \rangle. \quad (5.1)$$

This measurement was repeated for several pinhole locations.

Each plot in Figs. 5.2b-5.2d shows the results of repeatedly carrying out the computation in Eq. 5.1 for a reconstructor  $\mathbf{R}$  computed with each  $(\mathbf{r}, \boldsymbol{\theta})$  science

object location from  $10 < r < 30$  arc seconds in one arc second increments and  $0 < \theta < 2\pi$  radians in  $\pi/36$  radian increments. The red indicates the largest error and the blue the smallest error. The smallest error indicates the proper coordinate of the pinhole in the model space. Examining the figures, it can be seen that a reflection in the x-axis occurs between the model space co-ordinates and the system co-ordinates.

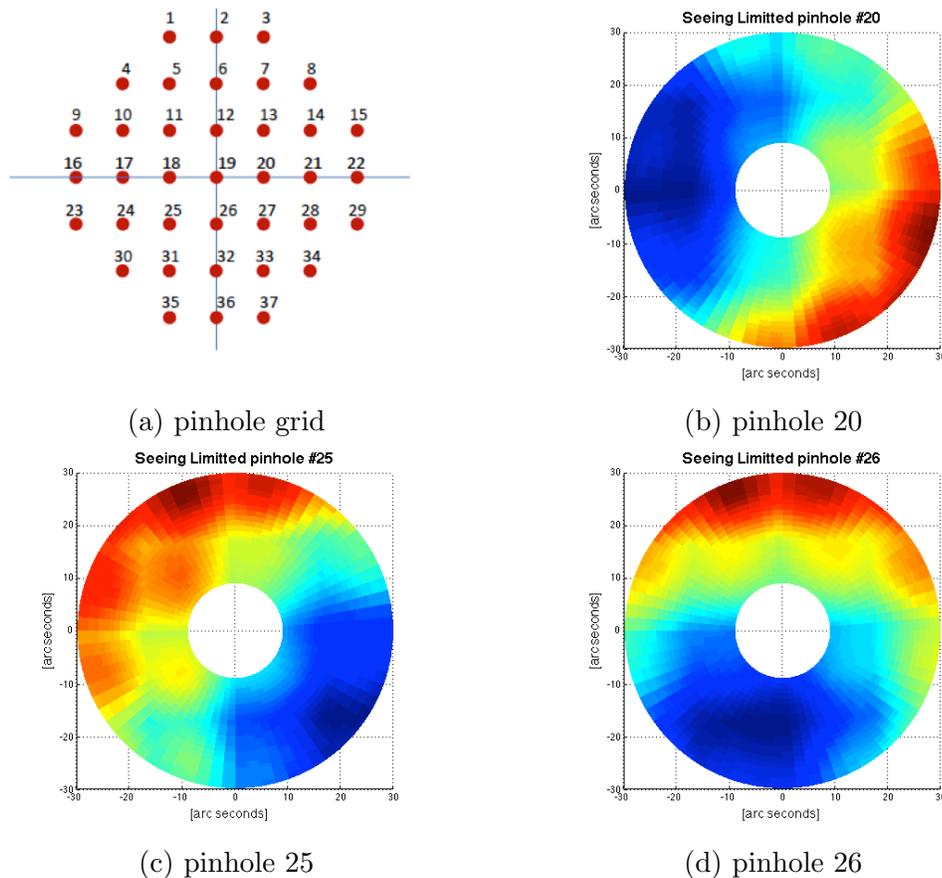


Figure 5.2: Measured locations of pinholes in Raven calibration unit relating model space coordinates to system coordinates.

### 5.1.2 CU neutral density filter magnitudes

The pinhole sources in the CU are fed by a single halogen lamp. The magnitude of the sources can be adjusted by varying the power of the lamp and/or by turning the filter wheel to one of 6 neutral density filters. The lamp does not provide uniform illumination across the pinhole grid therefore the magnitudes of the sources increase

with distance off-axis for any given setting. A plot of magnitude vs lamp power for each filter and a map of offsets is given in Fig. 5.3.

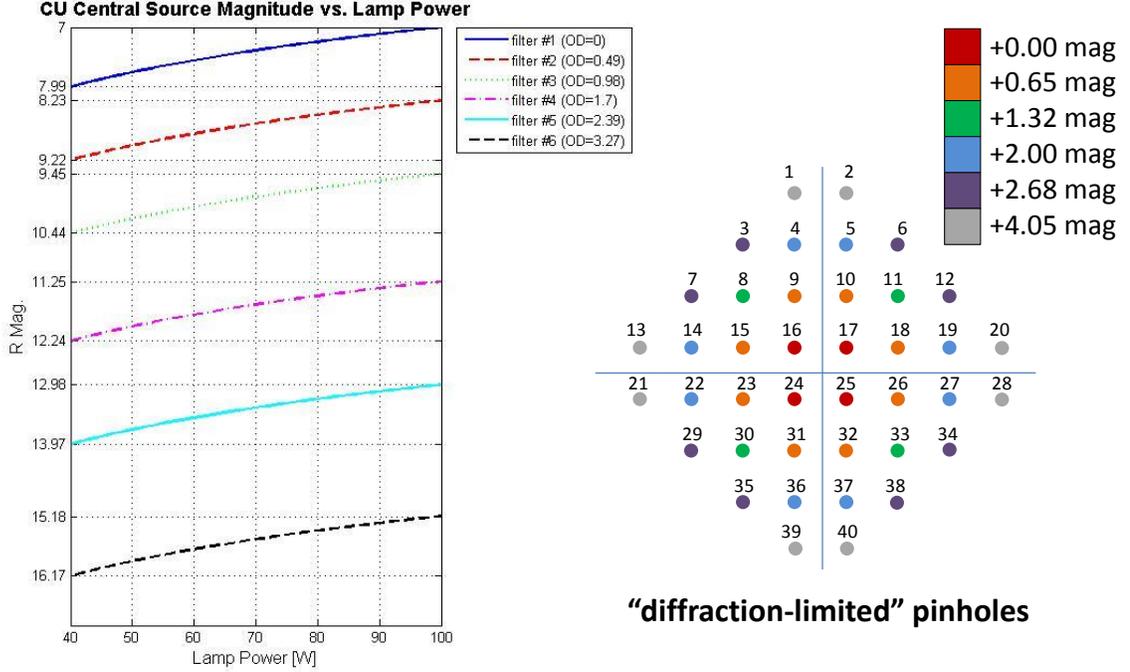


Figure 5.3: Source magnitudes provided by the CU as a function of pinhole location, filter and lamp power.

### 5.1.3 Transformation between WFS measurement bases

By design, the OL-WFSs of Raven are rotated with respect to each other and with respect to the CL-WFSs in the science arms (Fig. 5.4). The model-based tomographic reconstruction must be carried out in a common reference frame as the reconstructor makes no assumptions regarding the orientation of the WFS axes.

A transformation matrix is required to relate OL-WFS slopes in their local measurement space to slopes in a common measurement space [27],

$$\mathbf{s}'_{\alpha} = \Sigma_{(s'_{\alpha_i}, s_{\alpha_i})} \Sigma_{s_{\alpha_i}}^{-1} \mathbf{s}_{\alpha_i}. \quad (5.2)$$

Note that this does not transform the OL slopes measured in direction  $\alpha_i$  to slopes in direction  $\alpha'_i$  as a tomographic reconstructor would; it simply changes the basis onto which the slopes are projected in an optimal way using an MMSE solution. This is

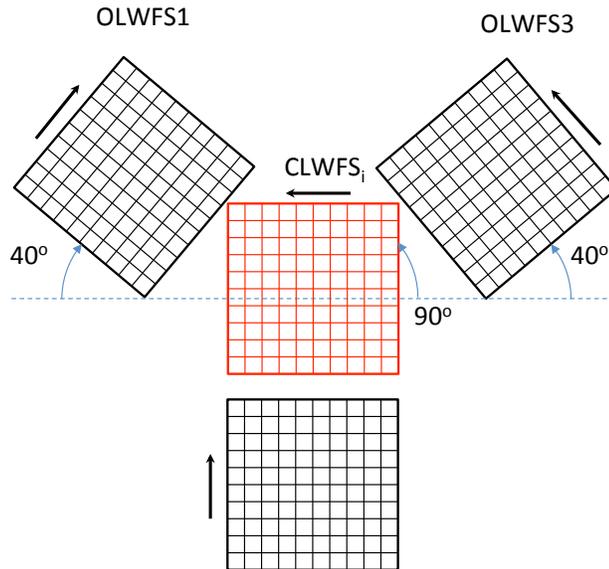


Figure 5.4: The OL-WFS axes are rotated with respect to each other and the CL-WFS axes.

a minor abuse of notation, as in this instance  $\alpha_i$  is not an angular direction, it is a reference to OL-WFS  $i$  where  $i = [1, 2, 3]$ . Similarly  $\beta_j$ ,  $j = [1, 2]$  indexes the CL-WFSs. It is convenient to choose the CL-WFS in each science arm as the reference frame and compute two separate reconstructors. This is convenient because the final step in the correction process is to apply the DM command matrix corresponding to each science arm and this is computed directly from the interaction matrix between the DM and the CL-WFS in each channel.

The components of this change of basis matrix can be computed from system measurements by applying a time series of common perturbations using the calibration DM. These are sensed by all WFSs simultaneously; the time series of slopes are labeled  $\mathbf{S}_{\alpha_i}$  from the OL-WFSs and  $\mathbf{S}_{\beta_j}$  from the CL-WFSs. The transformation matrix from OL-WFS  $i$  to CL-WFS  $j$  is therefore,

$$\mathbf{T}_{\alpha_i \rightarrow \beta_j} = \mathbf{S}_{\beta_j} \mathbf{S}_{\alpha_i}^T (\mathbf{S}_{\alpha_i} \mathbf{S}_{\alpha_i}^T)^{-1}. \quad (5.3)$$

Applying this transformation in Eq. 5.2 results in open-loop slopes expressed in the CL-WFS space, which can be rotated, translated and magnified with respect to the OL-WFS space. The expression is true provided the perturbations span the

vector space of the WFSs. In practice, the perturbations are introduced using a set of pre-computed CDM commands; two sets of perturbations were tested, the first was the Hadamard matrix of perturbations [95] and the second was a series of two dimensional sinusoidal waves. It was found that the second method provided slightly better results and was therefore integrated into the calibration pipeline. Later, a test using a precomputed set of CDM commands to play ground layer turbulence provided an even better result than the sine-wave sequence, however we must be cautious as this is the same ground layer turbulence that is played during tomographic tests.

In real-time operation, the OL-WFS slopes are pre-transformed to CL-WFS space and then fed to the tomographic reconstructor. Using Eq. 3.105 the slopes in the science direction are estimated and finally projected onto the DM influence-functions using a calibrated command matrix computed from the inverse closed-loop interaction matrix between CL-WFS and science DM using truncated singular value decomposition. In a single equation we have,

$$\mathbf{u}_\beta = \underbrace{(\mathbf{IM})^\dagger}_{\text{Command mat.}} \mathbf{R} \underbrace{\mathbf{T}_{\alpha_i \rightarrow \beta_j}}_{\text{Transformation}} \mathbf{s}_\alpha \quad (5.4)$$

A diagram of the calibration steps to measure the basis transformation matrices and the command matrices is given in Fig. 5.5

An alternative method which has been shown to provide similar performance quality is developed in [96]. The strategy of this method is to identify a set of meta-parameters which describe the WFS mis-alignments and generate a synthetic transformation matrix based on these parameters.

### Transforming between RTC and Matlab rasterization conventions

A simple, but important transformation of the WFS slope measurements is the one-to-one re-ordering matrix which takes the slopes read out by the RTC in a row-oriented rasterization, and transforms them to the column-oriented rasterization expected by the Matlab based reconstructor. By generating a series of simulated WFS frames and feeding these through both the RTC pixel processing and the simulation-based pixel processing, the slope ordering was compared and the centroiding accuracy assessed at the same time. The slope order transformation matrix is shown in Fig. 5.6, and the results of passing a simulated WFS images of Zernike mode 11 through both OOMAO and the RTC and then applying the transformation are shown in Fig 5.7.

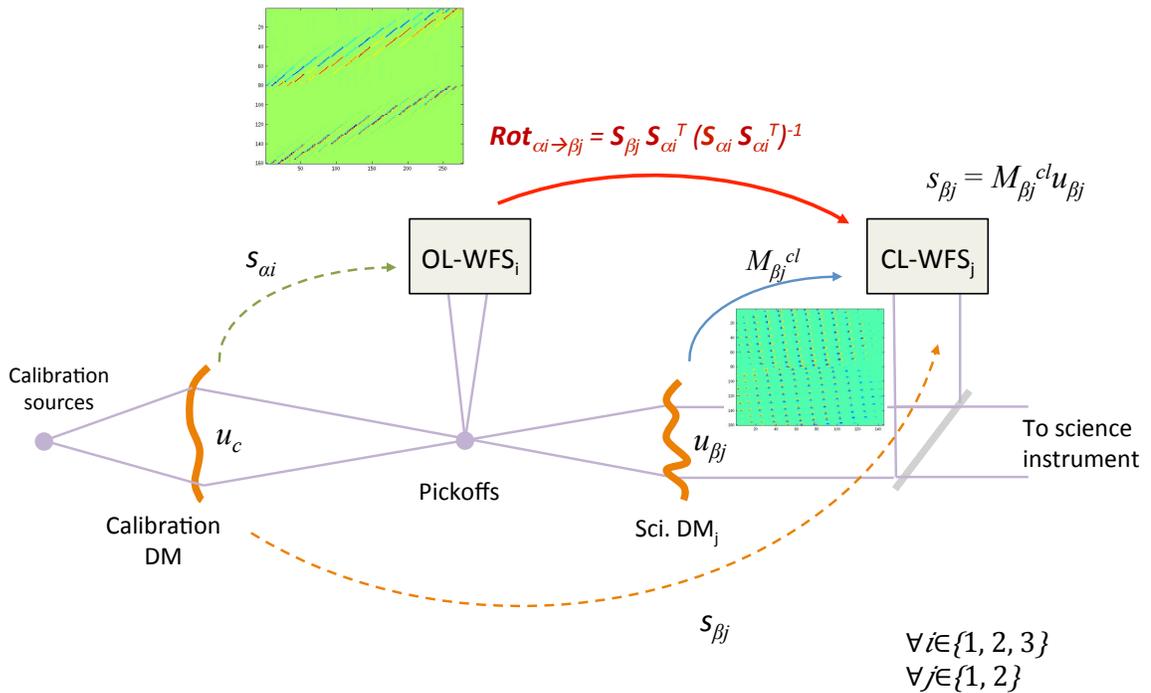


Figure 5.5: Common disturbances which span the vector space are applied to the ground-conjugated CDM and measured by all WFSs.

Small discrepancies are due to different levels of thresholding in the thresholded CoG spot centroiding algorithm which affects the result in the partially illuminated edge subapertures.

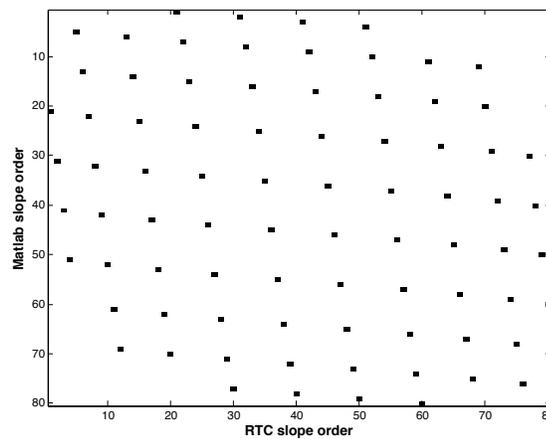


Figure 5.6: A one-to-one reordering matrix to link slopes readout of the RTC to computations carried out in Matlab.

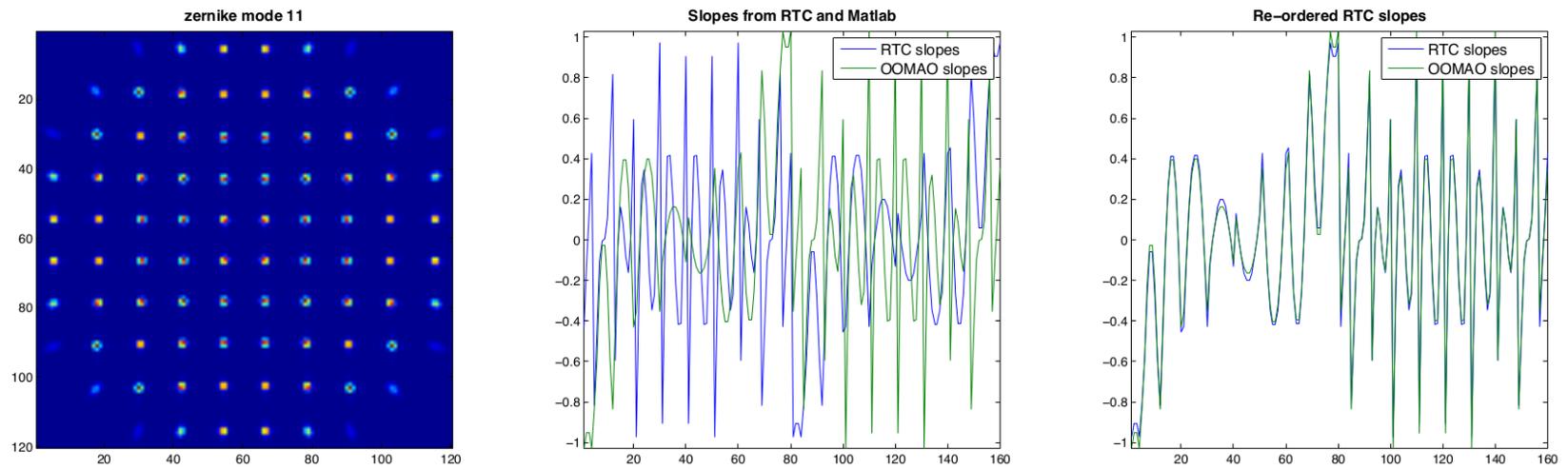


Figure 5.7: Left: simulated WFS frame of  $z_{11}$ . Middle: slopes measured by RTC and Matlab. Right: slopes measured by RTC and Matlab with order transformation matrix applied to RTC slopes

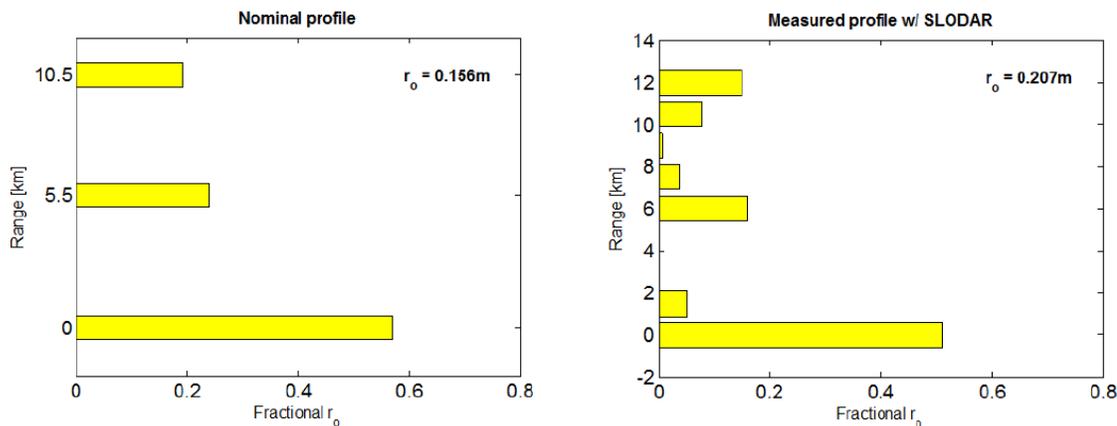
### 5.1.4 Identification of atmosphere parameters from the system

A critical element in tomographic WF estimation and temporal prediction is knowledge of the atmosphere. In the results presented here, it is assumed that the input atmosphere from the CU is well characterized. Some investigation into identification of the atmosphere, specifically the layer altitudes and strengths has been carried out by members of the Raven team. The on-board SLODAR method implemented on GeMS [11] was investigated as was the covariance fitting SLODAR implemented by the Canary team [27]. Some results from the lab and from the first on-sky run are shown in Fig. 5.8a and 5.8b

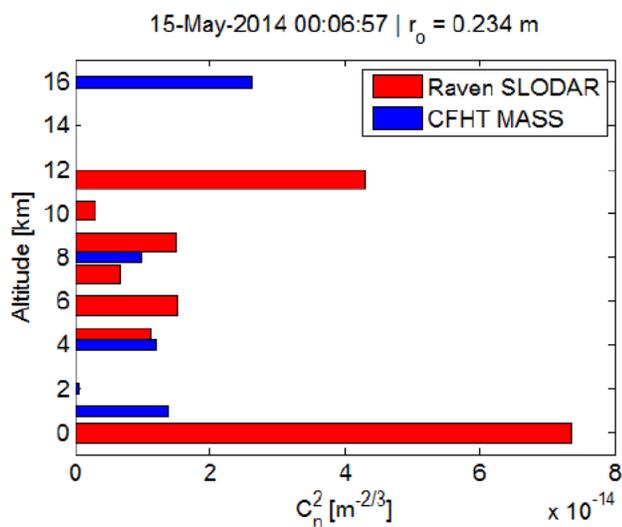
These estimates do not include wind velocity and direction. In the lab test results in this chapter, good knowledge of the wind profile is assumed based on the design of the CU phase screen rotation stages and CDM ground layer movie.

## 5.2 Static reconstructor results

In addition to confirming the results obtained in simulation, the purpose of testing the four types of static reconstructors (explicit-tomography and SA in both modal and zonal space) was two-fold. The first was to ensure Raven could perform wavefront correction and deliver the EE of 30% identified in the design specifications. The second was to establish a baseline from which the success or failure of both the simple predictive and LQG reconstructors could be deduced. It was possible to test the static explicit and AR-based modal reconstructors in the lab setting using the non-optimized Matlab code; however it will likely be excluded from on-sky experiments due to long computation times required to compute the reconstructors themselves. In the laboratory experiments, the atmosphere can be controlled and the same parameters maintained over the course of the testing. This is not true on-sky and the reconstructor must be updated at regular intervals by a background task receiving updated atmospheric parameters from a profiler such as the on-board SLODAR, mentioned in the previous section, or other available data. Due to the low order of Raven, parts of the code relating to background tasks such as reconstructor updating were not optimized for computation speed; therefore these methods are still of interest and could be tested on-sky given a commitment to optimized coding of the computations.



(a)



(b)

Figure 5.8: (a) Left: Expected CU atmosphere profile according to its design. Right: Measured CU atmosphere profile identified using on-board SLODAR. (b) Atmosphere identified during observation; results are compared to CFHT MASS measurements. There is no comparison for the ground layer as the CFHT data does not include a ground layer estimate. It would be different anyway due to Subaru dome seeing. The Raven profiler is constrained to estimate a maximum height of 12km and therefore attributes all turbulence from higher layers to that height.

### 5.2.1 Comparison of static MOAO methods

The first test to be carried out on the partially constructed system was the real world verification of the assertion made in Sec. 3.2.3 that the static SA and Explicit layered reconstructors are equivalent in that they produce the same final result by taking a different computational path (Figs 5.9). A single OL-WFS was used to move through the field and take measurements of a frozen turbulence frame at various positions corresponding to the pinhole grid. Using a post processing program we then selected certain configurations of guide star and science object asterisms and applied the static explicit and SA reconstructors in Zernike space, described in Sec. 3.3. The reconstructed phase in each of the science directions was then directly compared to the phase measured on the corresponding pinhole position by the WFS.

It was confirmed that explicit and SA methods produce the same results, as was shown in simulation. In the remaining tests, the Zonal SA static reconstructor is used, and subsequent predictive and LQG algorithms are based off of this method as it is the most computationally efficient of all options.

### 5.2.2 Learning on the fly

A method for characterizing the turbulence and generating a tomographic reconstructor based solely on system measurements is presented by the Canary team [27]. The method has been named Learn and Apply. In the lab and on engineering fields on-sky, this method can be tested using bright, compact science sources and has been shown to work very well on Raven in the lab (See sec. 5.2.3). It is used as a baseline, providing an estimate of the best performance the system can likely obtain while doing MOAO.

#### Hybrid Learn And Apply

In the case of a science observation, the necessary data from the CL-WFS directions will not be available, thus the first half of the reconstructor,  $\Sigma_{\beta,\alpha}$ , cannot be learned from the system during observation. A hybrid method is therefore being explored which takes advantage of all available measured data from the OL-WFSs during observation, and combines it with a model. The measured covariance matrix,  $\Sigma_{\alpha,\alpha}$  is shown in Fig.5.10.

Since the measured open loop covariance matrix,  $\Sigma_{\alpha,\alpha}$ , is available during obser-

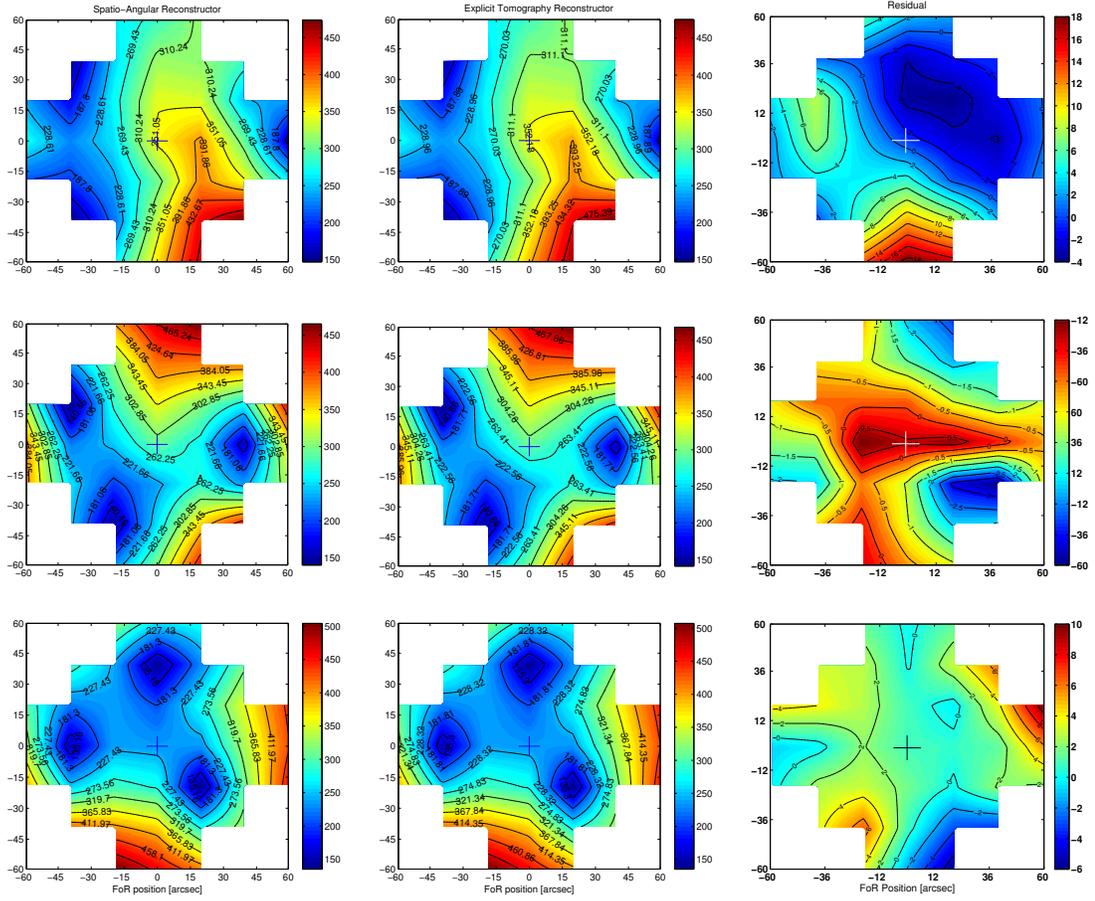
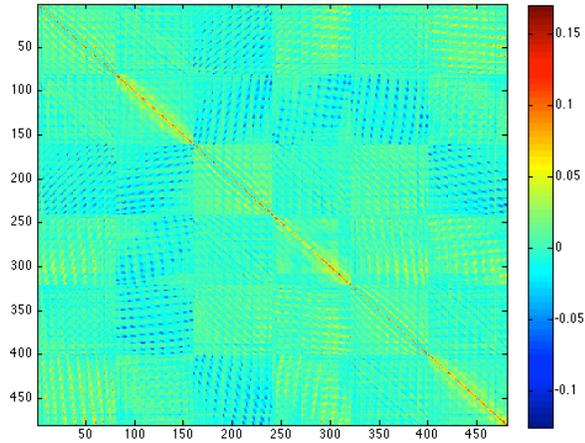
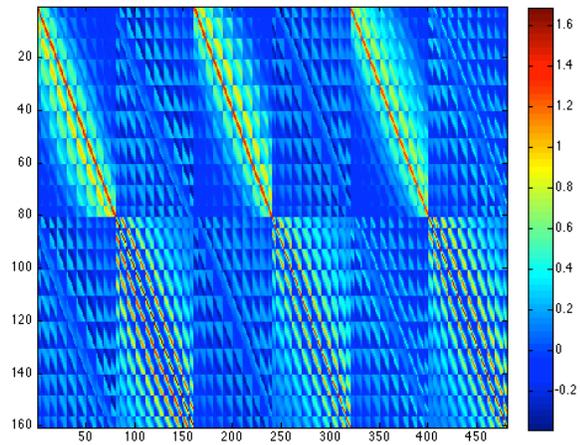


Figure 5.9: Comparison of residual wavefront error (in nm RMS) across the FoR with the spatio-angular (left column) and the explicit tomographic reconstructor (central column), for NGS asterisms 2', 1.5' and 1' apart. Differences plotted in the right column.

vation, we would like to use this data directly in the reconstructor. The challenge is to create a model of  $\Sigma_{\beta,\alpha}$  which matches that covariance measured by the system. The theoretical model built from an estimate of the  $C_n^2$  profile is shown in Fig. 5.11; this model makes the assumption that all WFSs are exactly aligned to the same axis, which is clearly not the case in general and in Raven specifically which by design has OL-WFSs rotated with respect to each other.

The model can be modified using the same calibration technique described in Sec. 5.1.3 which generates a matrix to transform slopes measured on one WFS to the space of another. In this case, the reverse transformation is required: the model is in the common reference frame and must be transformed into the individual OL-WFS

Figure 5.10: Measured  $\Sigma_{\alpha,\alpha}$ .Figure 5.11: modeled  $\Sigma_{\beta,\alpha}$ .

spaces using,

$$\mathbf{T}_{\beta_j \rightarrow \alpha_i} = \Sigma_{\alpha_i, \beta_j} (\Sigma_{\beta_j, \beta_j})^{-1}. \quad (5.5)$$

The expression for the model then becomes,

$$\hat{\Sigma}_{\beta, \alpha} = \langle \mathbf{s}_\beta (\mathbf{T}_{\beta \rightarrow \alpha} \mathbf{s}_\alpha)^T \rangle \quad (5.6)$$

$$= \langle \mathbf{s}_\beta \mathbf{s}_\alpha^T \mathbf{T}_{\beta \rightarrow \alpha}^T \rangle \quad (5.7)$$

$$= \Sigma_{\beta, \alpha}^{model} \mathbf{T}_{\beta \rightarrow \alpha}^T. \quad (5.8)$$

The transformed model is shown in Fig. 5.12a next to the measured covariance  $\Sigma_{\beta,\alpha}$  shown in Fig. 5.12b.

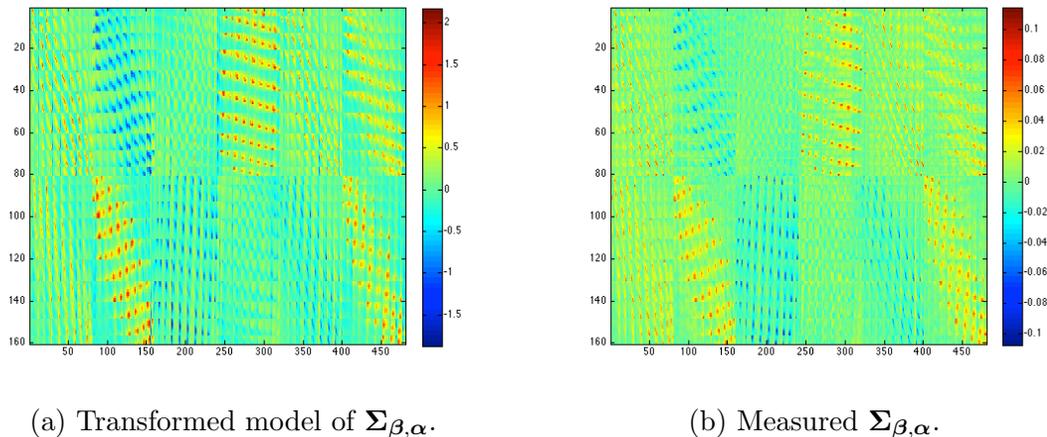


Figure 5.12: Comparison of a measured covariance matrix to a modelled covariance matrix.

Unfortunately, although there is reasonable visual agreement between the measured covariance matrix,  $\Sigma_{\beta,\alpha}$  and the transformed model of that matrix after scaling, there are obvious numerical differences between the two. The resulting reconstructor does not yet provide viable WF estimates, but the promise of being able to measure part of the reconstructor directly from the system holds high value and this method should continue to be explored.

### 5.2.3 Comparing MOAO to other WF correction modes

Multiple types of AO correction can be carried out with the Raven system; a comparison of the different modes provides an indication of the quality of correction being achieved with static MOAO compared to a simpler correction method, GLAO. The more ideal method which is unusable for science, Learn and Apply, and the best possible performance, SCAO are also implemented. In the first set of results (Fig. 5.13), a very ground layer dominated  $C_n^2$  profile was used due to an error in construction of the upper layer phase screens. In the second set of results (Fig. 5.14), the ground layer turbulence has been reduced to generate a more realistic atmosphere profile, however the total amount of turbulence is very low. In the final set of results, new phase screens with the correct physical parameters have been installed in the CU leading to a realistic  $C_n^2$  profile and amount of total turbulence. The atmospheric

parameters for the three sets of tests are summarized in Table 5.1.

Table 5.1: Atmospheric parameters in three test cases

Test case	$C_n^2$ profile	$r_0$
Fig. 5.13	0.90, 0.055, 0.045	0.27 m
Fig. 5.14	0.596, 0.224, 0.180	0.62 m
Fig. 5.15	0.596, 0.224, 0.180	0.156 m

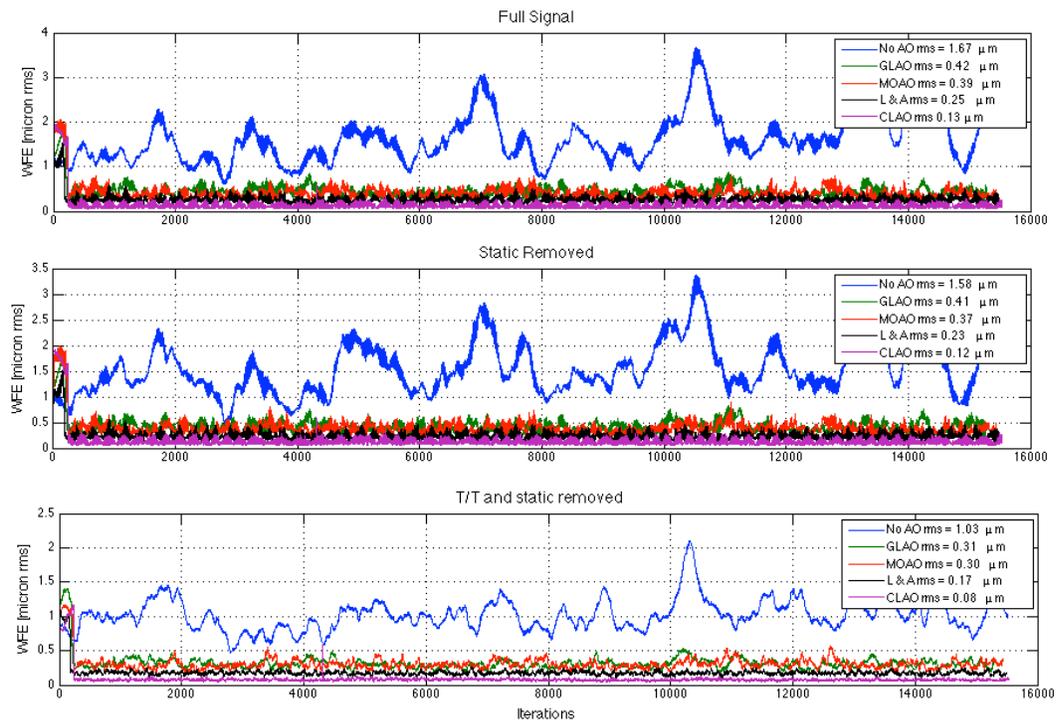


Figure 5.13: Static explicit zonal reconstruction compared to GLAO and SCAO in a ground layer-dominated case.

These results show that, even in a very ground layer dominated case, MOAO has the potential to improve over GLAO. In all cases the ideal Learn and Apply reconstructor out performs the model-based reconstructor, this is an indication that there are fundamental limitations in the model that do not fully capture either the system or the turbulence statistics.

Science images in J-Band were taken for a wide asterism case using both GLAO and MOAO with and without the LGS added on-axis. GLAO improves with the addition of the LGS, but not as much as MOAO. SCAO and no AO cases are included

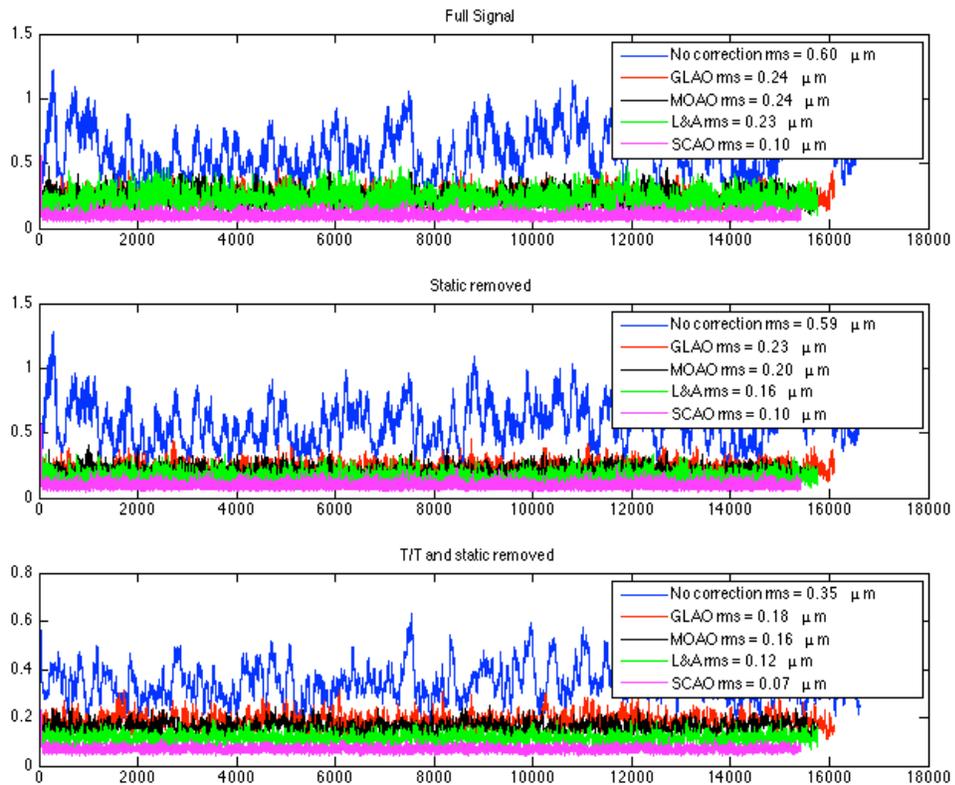


Figure 5.14: Static explicit zonal reconstruction compared to GLAO and SCAO in a low total turbulence case.

for comparison.

## 5.3 Dynamic reconstructor results

### 5.3.1 RTC implementation of the LQG

As stated in previous chapters, the stand-alone predictive model does not require any changes to the real time pipeline, the reconstructor can simply be replaced by the predictive matrix. The LQG however, requires a feedback loop in the RTC. To simplify the control pipeline within the RTC and to merge easily with less complex data pipeline cases, the control path outlined in Fig. 5.17a was manipulated such that it could be written as a simple feedback filter. Instead of applying the matrices  $\mathcal{A}_\delta^*$ ,  $\mathcal{M}_\infty$ ,  $\Gamma$  individually in the sequence shown,

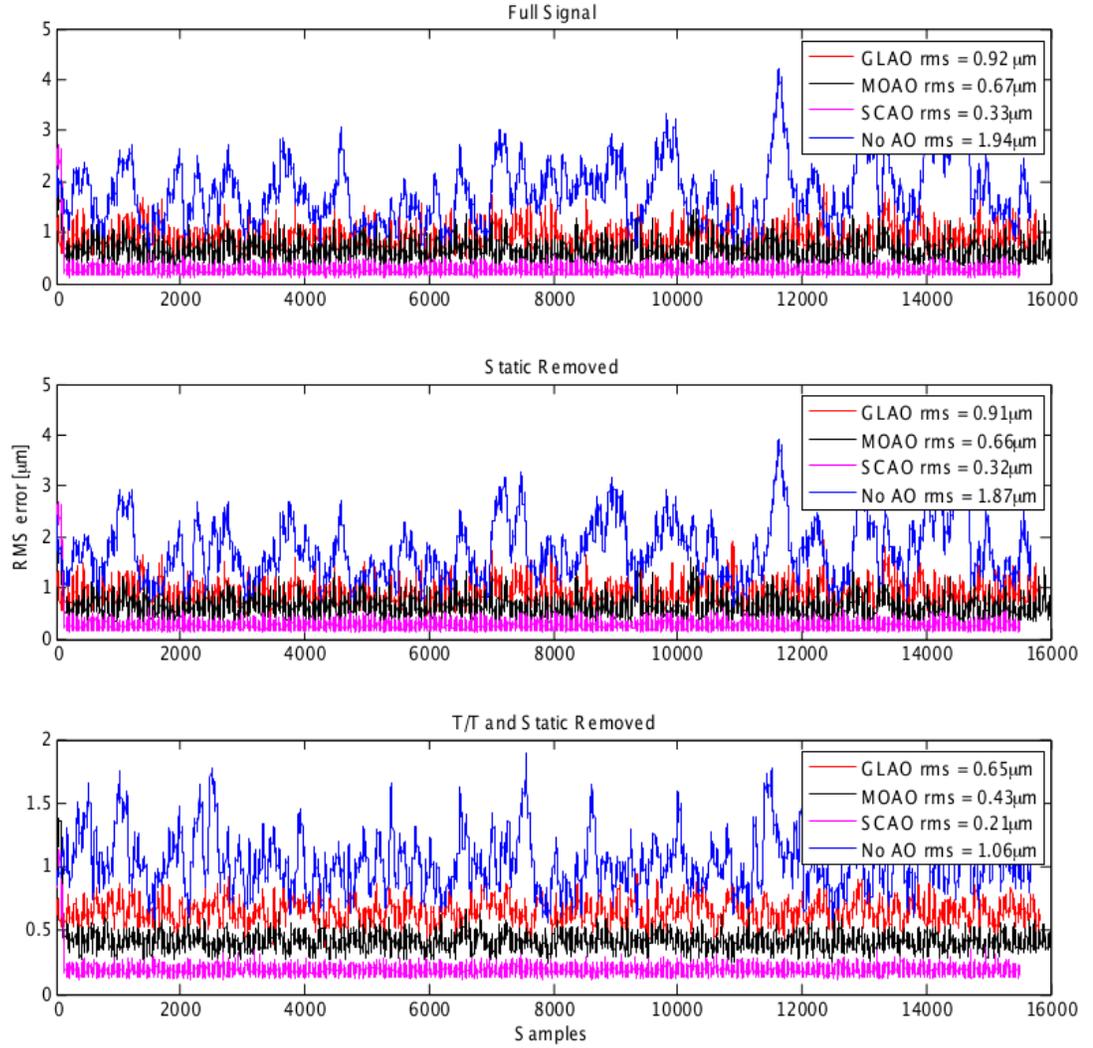


Figure 5.15: Static explicit zonal reconstruction compared to GLAO and SCAO in a realistic atmosphere case.

$$\begin{aligned}
 \hat{\mathbf{s}}_{k|k-1} &= \mathbf{\Gamma} \hat{\phi}_{k|k-1}, \\
 \hat{\phi}_{k|k} &= \hat{\phi}_{k|k-1} + \mathcal{M}_{\infty}(\mathbf{s}_k - \hat{\mathbf{s}}_{k|k-1}), \\
 \hat{\phi}_{k+1|k} &= \mathcal{A}_{\delta}^* \hat{\phi}_{k|k},
 \end{aligned} \tag{5.9}$$

the predicted phase,  $\hat{\phi}_{k+1|k}$ , can be expressed,

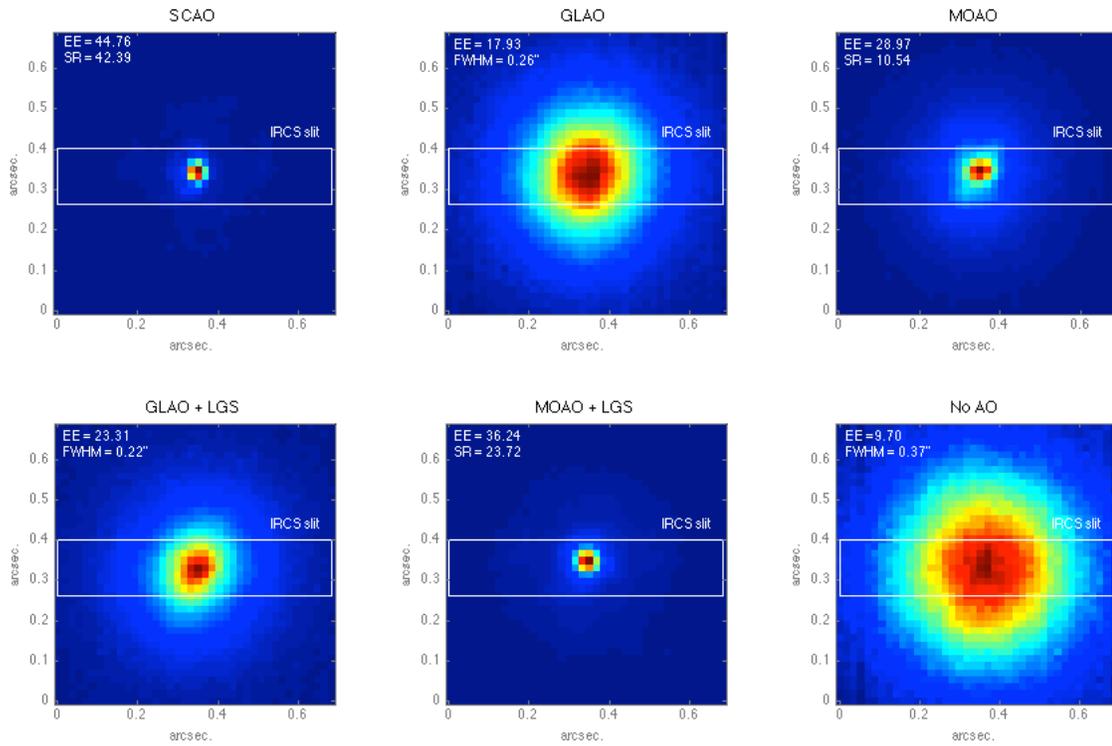


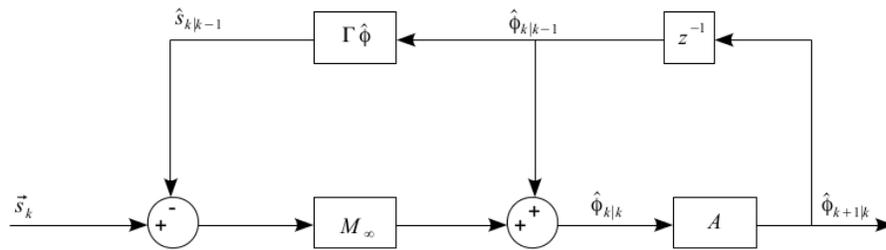
Figure 5.16: J-Band Science Images of GLAO vs MOAO for a static zonal SA tomographic reconstructor. Asterism diameter was 2 arc min and NGS magnitudes were 11.25, 12.13, and 12.13.

$$\begin{aligned}
 \hat{\phi}_{k+1|k} &= \mathcal{A}_\delta^* (\hat{\phi}_{k|k-1} + \mathcal{M}_\infty (\mathbf{s}_k - \Gamma \hat{\phi}_{k|k-1})) \\
 &= \mathcal{A}_\delta^* \hat{\phi}_{k|k-1} + \mathcal{A}_\delta^* \mathcal{M}_\infty \mathbf{s}_k - \mathcal{A}_\delta^* \mathcal{M}_\infty \Gamma \hat{\phi}_{k|k-1} \\
 &= \mathcal{A}_\delta^* \mathcal{M}_\infty \mathbf{s}_k + (\mathcal{A}_\delta^* - \mathcal{A}_\delta^* \mathcal{M}_\infty \Gamma) \hat{\phi}_{k|k-1}.
 \end{aligned} \tag{5.10}$$

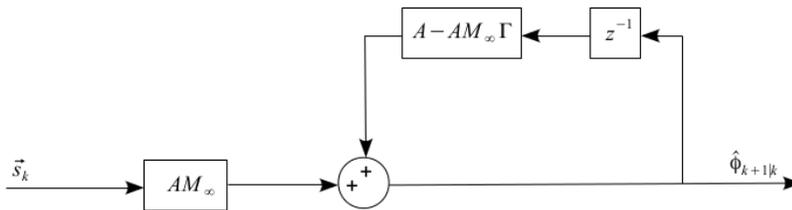
The control path can now be drawn as shown in Fig. 5.17b.

### 5.3.2 Results with WFS frame rates and guide star magnitudes

A cross section of measurements across NGS magnitude and WFS framerate were taken to show the improvement in performance using prediction and LQG control over the static reconstructor. These results were first obtained with the ground layer-dominated atmosphere profile (Fig. 5.19). As tomographic error is orthogonal to the



(a) LQG control scheme.



(b) LQG control scheme as implemented in the RTC.

Figure 5.17: (a) LQG control scheme. (b) LQG control scheme as implemented in the RTC.

temporal and noise error terms targeted by the prediction and LQG algorithms, it was predicted that the trends shown by these results would hold.

The data points in the following plots are labeled by reconstruction method and NGS magnitudes; the values 11.25, 13 and 15 are approximate values, the precise magnitudes of the GSs are found using Fig. 5.3 to be [10.8, 11.5, 11.5] for the data points labeled Mag 11.25, [12.6, 13.25, 13.25] for the data points labeled Mag 13, and [14.3, 15, 15] for the data points labeled Mag 15. The test asterisk is shown in Fig. 5.18.

Once the new phase screens were installed, the three algorithms were tested again to make sure that the improvement persisted under more challenging atmosphere conditions. The EE in a 140 milliarcsecond slit in the ground-layer dominated case was essentially 100% due to the low amount of turbulence and the Strehl ratios in the more turbulent atmosphere are very low so a direct comparison is difficult. However, it can be seen in Fig. 5.20 that the LQG algorithm is still very capable of improving results over the predictive and static reconstructors under realistic atmospheric conditions. The brightest data set is omitted. It shows limited improvement across the frequency space as there is already enough SNR at high frequencies; the LQG still improves over the other two algorithms, but there is little gain in reducing the

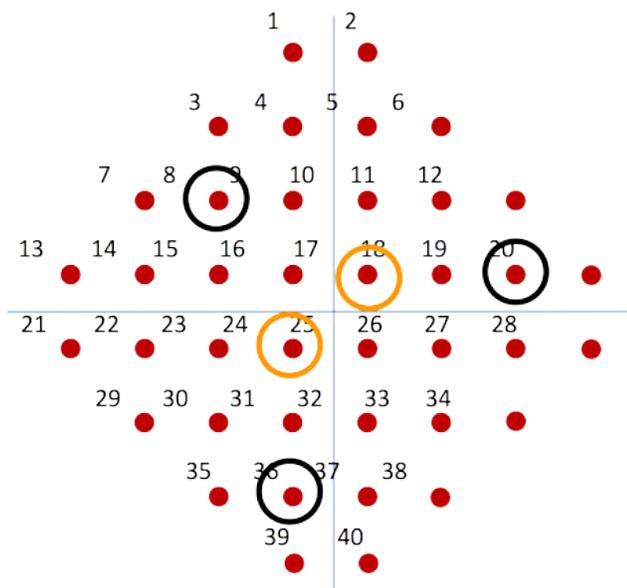


Figure 5.18: Asterism used in test cases.

framerate.

The performance achieved in lab tests is lower than that predicted by simulations in all reconstructor cases, but the trend of increasing peak performance by between 2 and 3 % with each increase in reconstructor complexity is reflected in both the simulation data and the measured data. The exception is the static reconstructor at magnitude 15 in the laboratory results; at this point, the signal on the WFSs becomes quite low and the thresholded centre of gravity (CoG) begins to fail in the lab setting. The overall decrease in performance can be attributed to multiple sources; these include imperfect calibration, underestimation of noise sources in the simulation compared to reality, effects of the rotated WFSs, DM fitting, OL go-to errors, and non-common path aberrations (NCPA) between the OL and science paths, as well as between the CL-WFSs and the science camera.

A trend noted in the laboratory measurements for brighter GSs is that the peak performance of the static and LQG reconstructors occur at rates slightly slower than the frequencies predicted by simulation, but the simple prediction performance peaks at a much slower frequency than expected. We speculate that this may be due to the  $r_0$  of the turbulence generated by the CU being a bit higher than expected; there may also be more noise than anticipated in the real system. As a result, the predictive algorithm would need to go to lower frequencies (greater temporal lag) to see the

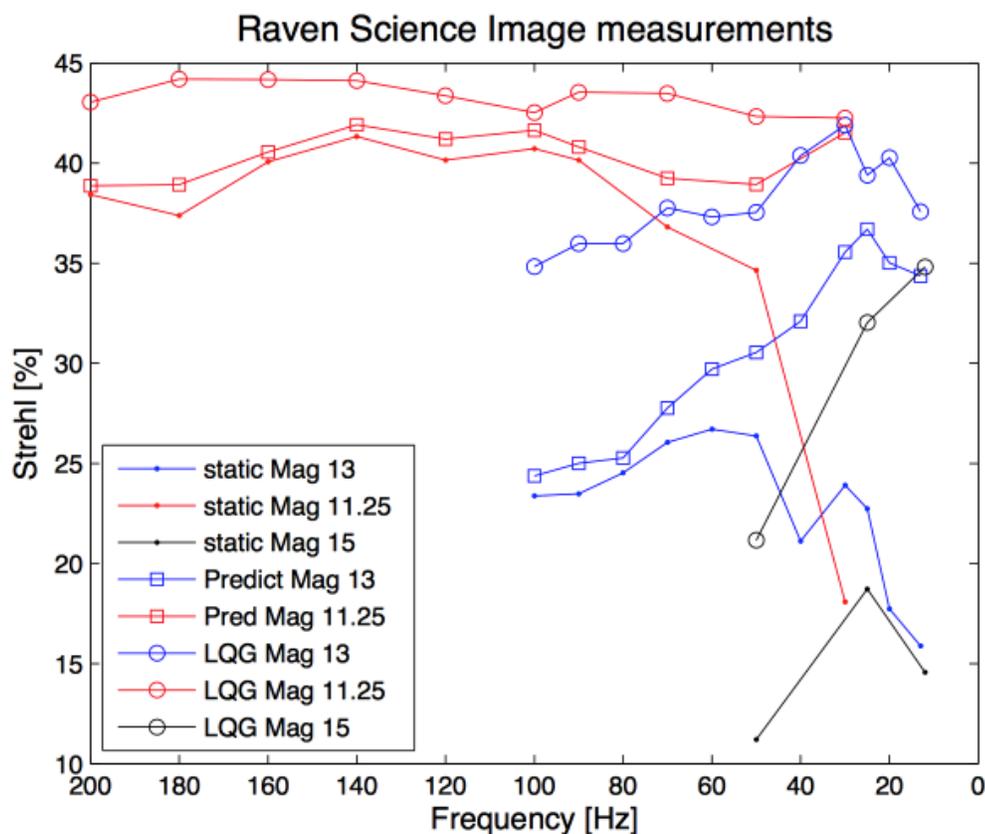


Figure 5.19: Results using static, predictive and LQG zonal reconstructors at varying framerate for different GS magnitudes using the ground layer dominated atmosphere.

most gains. The increased noise may also push the peak performance of the static reconstructor toward lower frequencies, although not as much because temporal lag errors are not addressed. Because the LQG handles both noise and temporal lag, the decrease in turbulence will still push the peak to slower frame rates, but an increase in noise will have less effect on the LQG than on the simple predictor. This trend is not reflected in the measurements using dimmer GSs, however the shifting of the peaks toward lower frequencies may mean that the system cannot be run slow enough to spot the new peaks in the data.

The science images in J-Band in Fig. 5.21 represent the best performance achieved for each reconstruction method for a fixed magnitude, they clearly show that the EE is increased and the spot image becomes progressively smaller for both predictive and LQG algorithms over the static reconstructor.

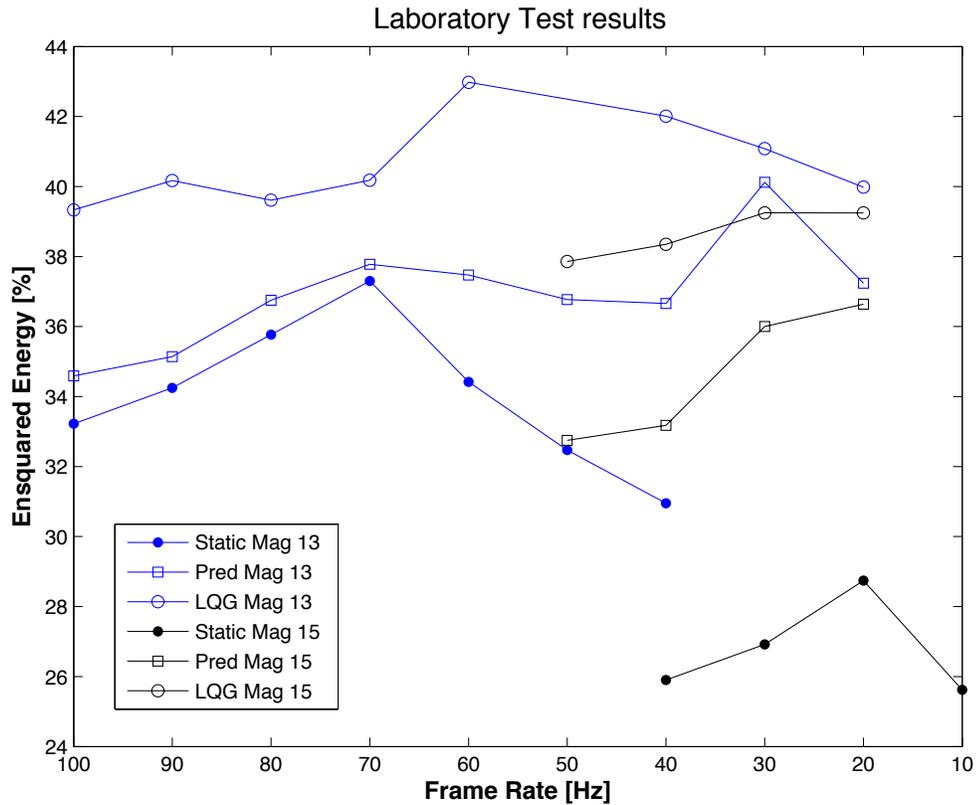


Figure 5.20: Performance comparison for Static, Predictive and LQG tomographic reconstruction algorithms with varying framerate using the distributed atmosphere profile.

## 5.4 Robustness to dynamic errors

There are several dynamic elements to observing on-sky that must be accounted for. In this section the results are presented from a preliminary investigation into the ability of the static reconstructor to continue to correct the WF under changing conditions. The reconstructor must be updated periodically to reflect these changing conditions, but how often this must be done is an open question and which parameters will drive the update period depends on the sensitivity of the system to errors as well as our ability to measure those parameters.

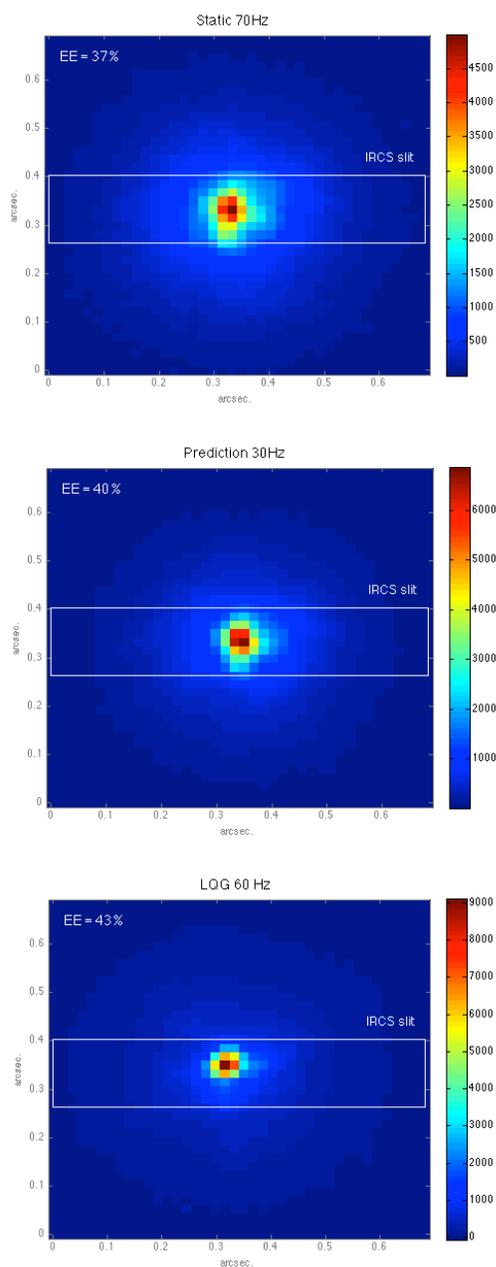


Figure 5.21: PSFs of best performance science images for the static, predictive and LQG algorithms.

#### 5.4.1 Sensitivity to atmospheric model layer altitude error

Running a static MOAO test with one atmospheric layer at a time and varying the input altitude of that layer in the model shows the sensitivity of the system to change in modelled layer altitude, as well as indicating either a possible discrepancy between

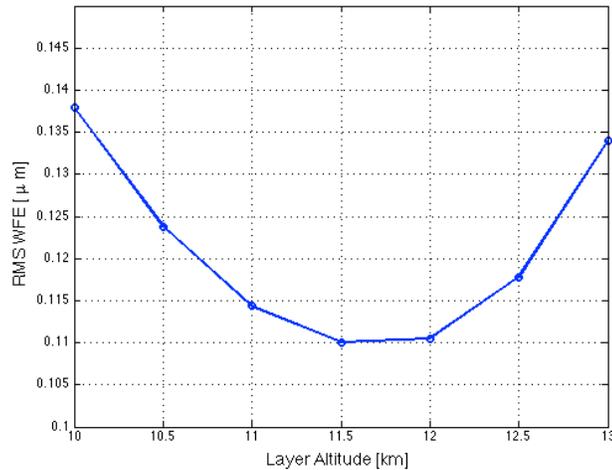


Figure 5.22: Variation of the single-layer atmosphere model used in the reconstructor shows the behaviour of the residual error as a function of model layer altitude.

model and system, or providing a more accurate characterization of the CU layer altitudes. The change in residual rms with model layer altitude for the 10km layer in Fig. 5.22 shows an obvious minimum at 11.5km. The results indicate an approximate increase in error in quadrature of 50% when the layer altitude is underestimated by 1km and assuming a close to linear trend, approximate doubling in quadrature when underestimated by 2km. The system appears slightly less sensitive to over estimation of altitude.

The full turbulence residual WFE is shown in Fig. 5.23 where the initial model contains an error and is then updated to something better. This demonstrates both the effect of the error and the functionality of the system to update smoothly on the fly without interrupting real-time correction.

It is becoming apparent that a tomographic system can be made less sensitive to layer altitude errors by estimating more layers; this was found in simulation (4.5.3) and in laboratory measurements where the use of the profile in the left side of Fig. 5.8a resulted in an rms error of  $0.43\mu m$  and the use of the measured profile in the right side of the figure resulted in an rms error of  $0.39\mu m$ , an improvement in quadrature of  $0.18\mu m$  rms, despite the fact that it is known that there are only three discrete input layers. Some of the improvement comes from the improved estimate of  $r_0$ , but just changing  $r_0$  without using the measured profile does not yield as much improvement. This result indicates that we may be able to make the reconstructor more robust against altitude drift between reconstructor updates by modelling a smoother  $C_n^2$

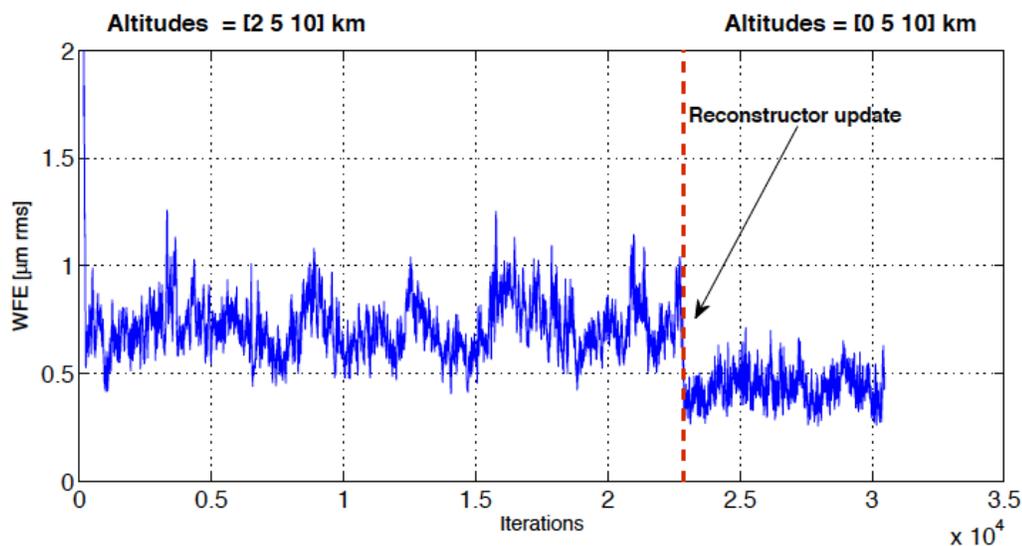


Figure 5.23: Update of the layer altitude profile during an observation.

profile.

#### 5.4.2 Field rotation update constraints

Another dynamic aspect to the system, beyond the changing over time of the atmosphere parameters, is the field rotation as the telescope tracks. This changes the relative positions of the GSs with respect to each other and is an issue specific to Raven due to its optical design and inability to use Subaru’s image derotator. To compensate for the field rotation, the reconstructor must be updated periodically with the new GS positions throughout an on-sky observation. The curves of the residual on CL-WFS1 in Fig. 5.24 show the effect of updating the reconstructor to a more accurate representation of current GS positions in the middle of the observation.

The error shown is due to a  $45^\circ$  rotation of the modelled GS positions, this is a very large error as the field rotation is very slow except near zenith where the maximum rotation speed will spike briefly to approximately  $3^\circ$  per minute, so although this parameter must be updated, it will likely not be a factor in driving the update rate of the reconstructor.

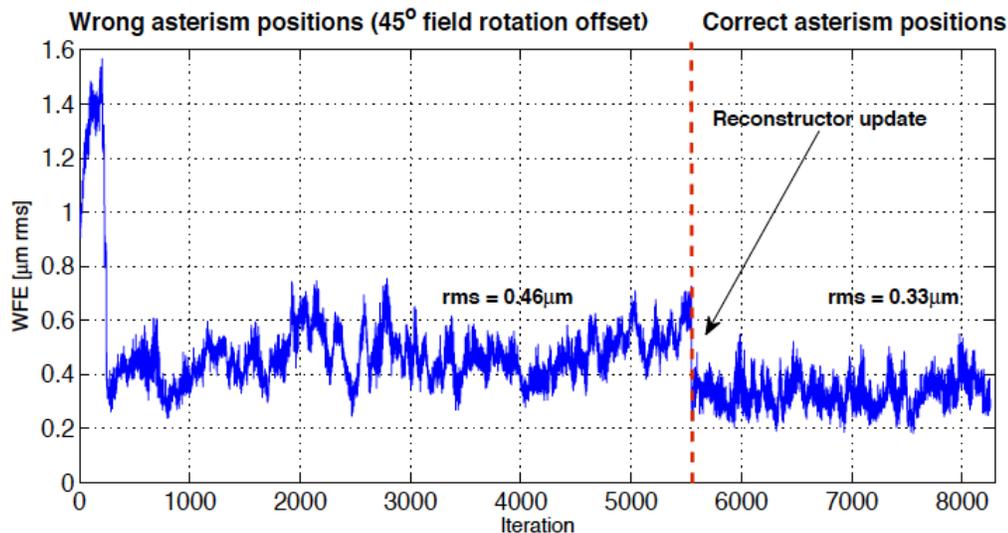


Figure 5.24: Update of the GS positions during an observation.

## 5.5 First on-sky results

Static MOAO reconstruction was successfully carried out on-sky using the Subaru telescope and the Raven science imager; the results are given in Fig. 5.25 and show a clear improvement using MOAO over no correction and GLAO in two science directions simultaneously. These results were obtained using the field shown in Fig. 5.26 where the OL-WFSs were on stars 5, 6 and 15, and the science pickoffs were on stars 11 and 7. The approximate diameter of this asterism is 2 arc-minutes and the magnitudes of the GSs are  $R = [12.8, 10.2, 12.4]$  respectively.

Raven was able to feed IRCS with MOAO corrected images. An example of this is seen in the images of Saturn shown in Fig. 5.27, where three of the moons of Saturn were used as NGSs and the science pickoffs were used to image two separate areas of the planet.

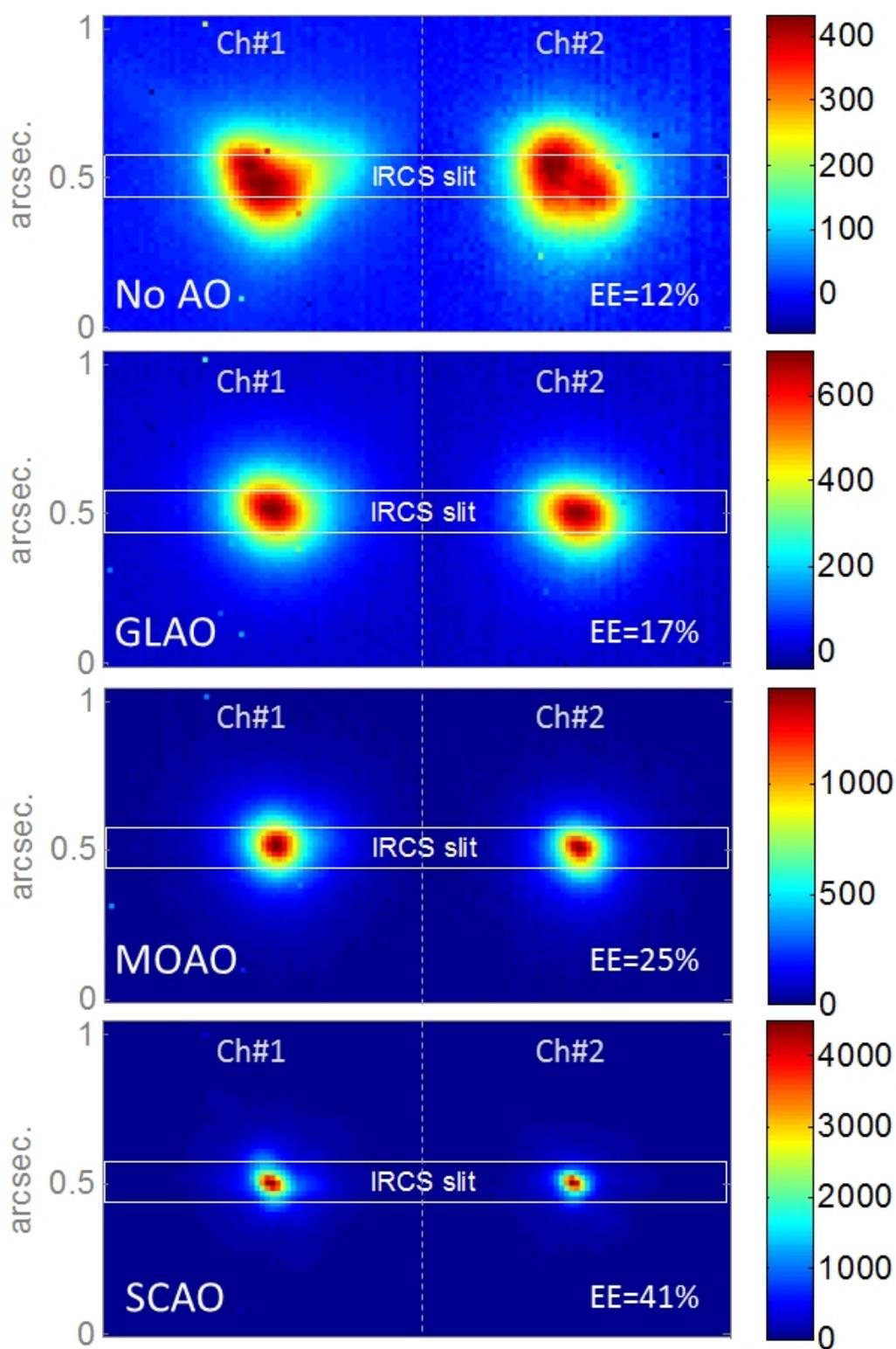


Figure 5.25: First on-sky results: Images from Raven science camera show performance for different AO modes [7].

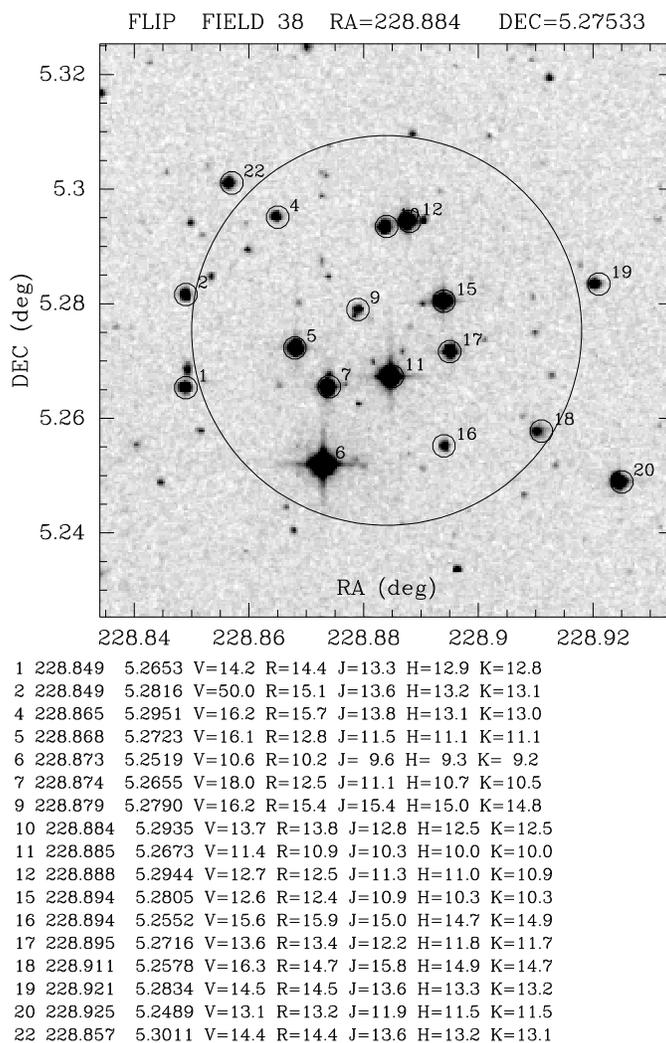


Figure 5.26: Observation field used for first on-sky MOAO test.

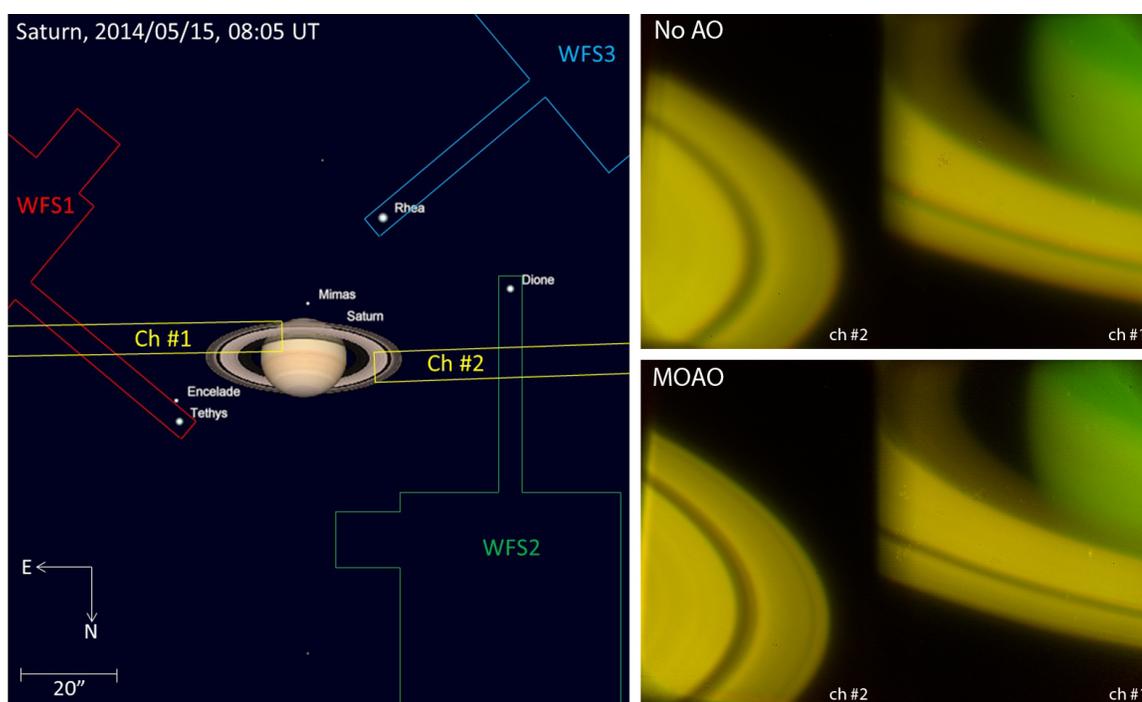


Figure 5.27: Saturn imaged on IRCS with and without MOAO correction [7].

# Chapter 6

## Conclusions

Monte-Carlo simulations of an MOAO system using static and dynamic tomographic reconstructors have been developed in both modal and zonal bases. The results have shown that a SA reconstructor, which derives no explicit information about the phase in the atmospheric layers, yields the same result as the more computationally complex reconstructor which does generate this information. As a result, the SA formulation has been extended to the more complex algorithms, the first of which uses temporal prediction alone and the second a spatio-temporally optimal LQG controller. This method has been identified as the appropriate technique to use in the specific case of an MOAO system which only requires an estimate of the phase in the pupil-plane.

Theory and simulation have shown that a gain in the limiting magnitude of Raven can be achieved by exploiting the trade-off between increasing lag error and improved SNR, by executing temporal prediction to reduce lag error. It was estimated that a gain of one magnitude can be expected from the use of a simple prediction within the tomographic step while reducing the overall framerate of the system. This can be achieved with no increase in the computational complexity of the real time pipeline. It was also estimated that a gain of two or more magnitudes can be expected if the LQG algorithm using one of the various prediction models presented in this dissertation is used; the most gain is achieved using the SA prediction models (whether in modal or zonal space). It is concluded that the best algorithm is the zonal SA LQG formulation which allows the most minimal representation of the phase and makes the prediction directly in the pupil-plane. There is an increase in the number of computations over that static case, but it is a much more conservative increase compared to the explicit layered LQG algorithm which, on a system of the scale of Raven, requires a factor of 20 more real-time computations per iteration.

Static reconstructors, implemented in simulation, have been successfully ported to the Raven MOAO test bed. Experimental testing has also validated the simulations and shown that a pupil-plane WF reconstructor is sufficient for an MOAO system and equal to the more computationally intensive explicit layered estimators. The static SA zonal reconstructor has performed on-sky tomography and delivered EEs close to the target of at least 30% EE simultaneously for two science targets. It has been shown definitively that model-based MOAO out-performs GLAO in the laboratory and on-sky.

The more complex dynamic reconstructors in a zonal basis have been implemented on the Raven system; the simulation results are supported by laboratory measurements taken on Raven with the telescope simulator. They show that the LQG algorithm delivers higher EE than the SA prediction algorithm which in turn improves over the static reconstructor. A comparison of performance for various GS magnitudes has shown that the limiting magnitude of Raven can be improved by two or more magnitudes by using the LQG controller provided knowledge of the wind profile and other atmospheric parameters are known to a high enough accuracy. A preliminary investigation of robustness indicates that it is sufficient to have an estimate of the wind speed of any layer within a factor of 2 of the actual value and the wind direction within  $+/- 50^\circ$  of the actual direction.

The cumulative result of the development and comparison of all the static and dynamic tomographic reconstructors in this chapter is a clear understanding of the process of doing tomography for MOAO and the specific properties that can be taken advantage of to simplify it as much as possible.

## 6.1 Future Work

The dynamic algorithms developed in this dissertation and validated through simulation and in the lab will be tested on-sky. Based on the quality of these on-sky results, it would be advisable to take advantage of Raven's telescope and atmosphere simulation abilities to do a broad parameter space study to test the robustness and the stability of the LQG controller to various errors under well known conditions. This will help to set error budgets for current and future atmospheric profilers and prioritize the parameters which have a large effect on performance over those that have less effect.

The identification of vibrations in the system, while Raven is on the Naysmith

platform of Subaru telescope, along with the addition of corresponding vibration filters in the LQG controller is a task that is also planned for the future.

# Bibliography

- [1] Valley, G. C., “Isoplanatic degradation of tilt correction and short-term imaging systems,” *Applied Optics* **19**(4), 574–577 (1980).
- [2] Masciadri, E., Stoesz, J., Hagelin, S., and Lascaux, F., “Mt. Graham: optical turbulence vertical distribution with standard and high resolution,” *Proc. SPIE* **7733**, 1P (2010).
- [3] Roggemann, M. C. et al., [*Imaging Through Turbulence*], Springer (1996).
- [4] Tokovinin, A., “Adaptive optics tutorial at ctio.” <http://www.ctio.noao.edu/~atokovin/tutorial/part3/wfs.html> (2001). [Online; accessed Feb. 2014].
- [5] Sacek, V., “Telescope aberrations: types and causes.”
- [6] Noll, R. J., “Zernike polynomials and atmospheric turbulence,” *Opt. Soc. Am.* **66** (1976).
- [7] Lardière, O., Andersen, D. R., et al., “MOAO results with RAVEN,” *Proc. SPIE* **9148**, 50 (2014).
- [8] Nelson, J., “Thirty-Meter Telescope: progress and approaches,” *Proc. SPIE* **6986**, 02 (2008).
- [9] Ramsay, S., D’Odorico, S., Casali, M., Gonzalez, J. C., Hubin, N., Kasper, M., Kuffl, H. U., Kissler-Patig, M., Marchetti, E., Paufique, J., Pasquini, L., Siebenmorgen, R., Richichi, A., Vernet, J., and Zerbi, F. M., “An overview of the E-ELT instrumentation programme,” *Proc. SPIE* **7735**, 24 (2010).
- [10] Shectman, S. and Johns, M., “GMT overview,” *Proc. SPIE* **7733**, 1Y (2010).

- [11] Cortés, A., Neichel, B., Guesalaga, A., Osborn, J., Rigaut, F., and Guzman, D., “Atmospheric turbulence profiling using multiple laser star wavefront sensors,” *Monthly Notices of the RAS* **427**(3) (2012).
- [12] Tallon, M. and Foy, R., “Adaptive telescope with laser probe: isoplanatism and cone effect,” *A & A* **235** (1990).
- [13] Tokovinin, A. et al., “Optimized modal tomography in adaptive optics,” *Astronomy and Astrophysics* **378** (2001).
- [14] Johnston, D. C. and Welsh, B. M., “Atmospheric turbulence sensing for a multiconjugate adaptive optics system,” *Proc. SPIE* **1542**, 76–87 (1991).
- [15] Ragazzoni, R., Marchetti, E., and Rigaut, F., “Modal tomography for adaptive optics,” *A&A* **L53** (1999).
- [16] R. Flicker, F. J. Rigaut, B. L. E., “Comparison of multiconjugate adaptive optics configurations and control algorithms for the gemini south 8-m telescope,” *Proc. SPIE* **4007**, 1032–1043 (2000).
- [17] Rigaut, F., Ellerbroek, B., and Flicker, R., “Principles, limitations and performance of multiconjugate adaptive optics,” *Proc. SPIE* **4007**, 1022–1031 (2000).
- [18] Hammer, F. et al., “The FALCON concept: Multi-Object spectroscopy combined with MCAO in near-IR,” in [*Scientific Drivers for ESO Future VLT/VLTI Instrumentation*], Bergeron, J. and Monnet, G., eds., Springer (2002).
- [19] Puech et al., “Falcon: Extending adaptive correction to cosmological fields,” *New Ast. Rev.* **50** (2006).
- [20] Primmerman, C. A. et al., “Compensation of atmospheric optical distortion using a synthetic beacon,” *Nature* **353** (1991).
- [21] Ellerbroek, B. L., “Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques,” *J. Opt. Soc. Am* **19** (2002).
- [22] Gavel, D. et al., “Visible light laser guidestar experimental system (ViLLaGEs): on-sky tests of new technologies for visible wavelength all-sky coverage adaptive optics systems,” *Proc. SPIE* **7015**, 70150G (2008).

- [23] Wizinowich, P. et al., “W. M. Keck observatory’s next-generation adaptive optics facility,” *Proc. SPIE* **7736**, 0K (2010).
- [24] Ammons, S. M. et al., “Integrated laboratory demonstrations of multi-object adaptive optics on a simulated 10 meter telescope at visible wavelengths,” *PASP* **122**, 573–589 (2010).
- [25] Andersen, D. R., Fischer, M., and Véran, J.-P., “Building an open loop interaction matrix for VOLT,” in [*OSA Optics & Photonics Technical Digest*], A0–A4, Optical Society of America (2009).
- [26] Andersen, D. R. et al., “VOLT: the victoria open loop testbed,” *Proc. SPIE* **7015**, 70150H (2008).
- [27] Vidal, F., Gendron, E., and Rousset, G., “Tomography approach for multi-object adaptive optics,” *J. Opt. Soc. Am. A* **27**, A253–A264 (Nov 2010).
- [28] Gendron, E. et al., “Status update of the CANARY on-sky MOAO demonstrator,” *Proc. SPIE* **7736**, 77360P (2010).
- [29] Morris, T. et al., “CANARY: The NGS/LGS MOAO demonstrator for EAGLE,” in [*the 1st AO4ELT Conference - Adaptive Optics for Extremely Large Telescopes proceedings*], EDP Sciences (2010).
- [30] Gendron, E. et al., “MOAO first on-sky demonstration with CANARY,” *A&A* **529** (2011).
- [31] Vidal, F., Gendron, E., et al., “Analysis of on-sky MOAO performance of CANARY using natural guide stars,” *A & A* (2014).
- [32] Morris, T. J., Gendron, E., et al., “CANARY phase B: on-sky open-loop tomographic LGS AO results,” *Proc. SPIE* **9148**, 52 (2014).
- [33] Tokunaga, A. T. et al., “Infrared camera and spectrograph for the subaru telescope,” *Proc. SPIE* **3354**, 512–524 (1998).
- [34] Kolmogorov, D. N., “Dissipation of energy in the locally isotropic turbulence,” in [*Proc. of the USSR Academy of Sciences*], **32** (1941).
- [35] Tatarski, V. I., [*Wave propagation in a turbulent medium (R. A. Silverman, trans.)*], McGraw-Hill book company, Inc. (1961).

- [36] Khintchine, A., “Kerrelationstheorie der stationären stochstischen prozesse,” *Mathematische Annalen* **109**, 604–615 (1934).
- [37] Tatarski, V., [*The effects of the turbulent atmosphere on wave propagation*], Israel Program for Scientific Translations (1971).
- [38] Obukhov, A. M., “Structure of the temperature field in a turbulent flow (U.S.A. department of the army, trans.),” *Izvestiia Akademii Nauk U.S.S.R., Ser. Geogr. I Goofiz.* **2**(1), 58–69 (1949).
- [39] Yaglom, A. M., “On the local structure of the temperature field in a turbulent flow,” *Notes of the Academy of Sciences, USSR* **69**(6) (1949).
- [40] Hufnagel, R. E. and Stanley, N., “Modulation transfer function associated with image transmission through turbulent media,” *JOSA* **54**(1), 52–60 (1964).
- [41] Tyson, R. K., [*Principles of Adaptive Optics*], Academic Press, 2nd ed. (1998).
- [42] Roddier, F., [*Adaptive Optics in Astronomy*], Cambridge University Press (1999).
- [43] Fried, D. L., “Statistics of a geometric representation of wavefront distortion,” *JOSA* **55**(11), 1427–1431 (1965).
- [44] Madec, P.-Y., “Overview of deformable mirror technologies for adaptive optics and astronomy,” *Proc. SPIE* **8447**, 05 (2012).
- [45] Valley, G. C., “Long- and short-term strehl ratios for turbulence with finite inner and outer scales,” *Appl. Opt.* **18**, 984–987 (Apr 1979).
- [46] Conan, R., *Modelisation des effets de l’echelle externe de coherence spatiale du front d’onde pur l’observation a Haute Resolution Angulaire en Astronomie*, PhD thesis, Universite de Nice (2000).
- [47] Rigaut, F. and Neichel, B., “First on-sky results with GeMS,” *AO4ELT Conf* **2**, 2 (2011).
- [48] Andersen, D. R., Lardiere, O., Veran, J., Bradley, C., and Kerley, D. A., “Comparing the performance of open loop centroiding techniques in the Raven MOAO system,” *Proc. SPIE* **9148**, 216 (2014).

- [49] Gilles, L. and Ellerbroek, B., “Shack-hartmann wavefront sensing with elongated sodium laser beacons: centroiding versus matched filtering,” *Appl. Opt.* **45**(25), 6568–6576 (2006).
- [50] Keskin, O., Hampton, P., Conan, R., Bradley, C., Hilton, A., and Blain, C., “Woofers-tweeters adaptive optics test bench,” *Adaptive Hardware and Systems, NASA/ESA Conference* **1**, 74–80 (2006).
- [51] Andersen, D. R., Jackson, K. J., Blain, C., Bradley, C., Correia, C., Ito, M., Lardièrre, O., and Véran, J.-P., “Performance modeling for the raven multi-object adaptive optics demonstrator,” *PASP* **124**(915), pp. 469–484 (2012).
- [52] Andersen, D. R. et al., “Status of the Raven MOAO science demonstrator,” *Proc. SPIE* **8447**, 3F (2012).
- [53] Lardièrre, O. et al., “Final optical design of Raven: a MOAO science demonstrator for subaru,” *Proc. SPIE* **8447**, 53 (2012).
- [54] Lavigne, J.-F. et al., “Design and test results of the calibration unit for the MOAO demonstrator RAVEN,” *Proc. SPIE* **8447**, 54 (2012).
- [55] Pham, L. et al., “Raven calibration,” *Proc. Adaptive Optics for ELTs* **2** (2011).
- [56] TMT-Science-Advisory-Committee, “Thirty meter telescope detailed science case: 2007,” (2007). Tech. rep.
- [57] Robin, A. C., Reylé, C., Derrière, S., and Picaud, S., “A synthetic view on structure and evolution of the milky way,” *Astronomy&Astrophysics* **409** (2003).
- [58] Correia, C. et al., “Static and predictive tomographic reconstruction for wide-field multi-object adaptive optics systems,” *J. Opt. Soc. Am. A* **31** (2013).
- [59] Fusco, T. et al., “Optimal wave-front reconstruction strategies for multiconjugate adaptive optics,” *J. Opt. Soc. Am. A* **18** (2001).
- [60] Piatrou, P. and Roggemann, M. C., “Performance study of kalman filter controller for multiconjugate adaptive optics,” *Appl. Opt.* **46** (2007).
- [61] Anderson, B. D. O. and Moore, J. B., [*Optimal Filtering*], Dover Publications Inc. (2005).

- [62] Whiteley, M., Welsh, B. M., and Roggemann, M. C., “Optimal modal wavefront compensation for anisoplanatism in adaptive optics,” *J. Opt. Soc. Am. A* **15**, 2097–2106 (1998).
- [63] von Kármán, T., “Progress in the statistical theory of turbulence,” in [*Presented at the Heat Transfer and Fluid Mechanics Institute, Los Angeles, California*], (1948).
- [64] Vogel, C. R., “Sparse matrix methods for wavefront reconstruction revisited,” *Proc. SPIE* **5490**, 1327–1335 (2004).
- [65] Anderson, B. D. O. and Moore, J. B., [*Optimal Control, Linear Quadratic Methods*], Dover Publications Inc. (2005).
- [66] Le Roux, B., Conan, J.-M., Kulcsár, C., Raynaud, H.-F., Mugnier, L. M., and Fusco, T., “Optimal control law for classical and multiconjugate adaptive optics,” *J. Opt. Soc. Am. A* **21**, 1261–1276 (2004).
- [67] Petit, C. et al., “First laboratory validation of vibration filtering with LQG control law for adaptive optics,” *Opt. Exp.* **16**(1) (2008).
- [68] Correia, C., Véarn, J.-P., et al., “Increased sky coverage with optimal correction of tip and tilt-anisoplanatism modes in laser-guide-star multi conjugate adaptive optics,” *J. Opt. Soc. Am. A* **30**(4) (2013).
- [69] Chassat, F., “Calcul du domaine d’isoplanetisme d’un systeme d’optique adaptative fonctionnant a travers la turbulence atmospherique,” *J. Optics (Paris)* **20**(1), 13–23 (1989).
- [70] Takato, N. and Yamaguchi, I., “Spatial correlation of zernike phase-expansion coefficients for atmospheric turbulence with finite outer scale,” *J. Opt. Soc. Am. A* **12**, 958–963 (May 1995).
- [71] Whiteley, M. R., Roggemann, M. C., and Welsh, B. M., “Temporal properties of the zernike expansion coefficients of turbulence-induced phase aberrations for aperture and source motion,” *J. Opt. Soc. Am. A* **15**, 993–1005 (1998).
- [72] Winker, D. M., “Effect of a finite outer scale on the zernike decomposition of atmospheric optical turbulence,” *J. Opt. Soc. Am. A* **8**, 1568–1573 (Oct 1991).

- [73] Lundström, L. and Unsbo, P., “Transformation of zernike coefficients: scaled, translated, and rotated wavefronts with circular and elliptical pupils,” *J. Opt. Soc. Am. A* **24**(3), 569–577 (2007).
- [74] Correia, C., Raynaud, H.-F., Kulcsár, C., and Conan, J.-M., “On the optimal reconstruction and control of adaptive optical systems with mirror dynamics,” *J. Opt. Soc. Am. A* **27**, 333–349 (Feb. 2010).
- [75] Wiener, N., “Generalized harmonic analysis,” *Acta Mathematica* **55**, 117–258 (1930).
- [76] Conan, J.-M., Rousset, G., and Madec, P.-Y., “Wave-front temporal spectra in high-resolution imaging through turbulence,” *J. Opt. Soc. Am. A* **12**, 1559–1570 (1995).
- [77] Taylor, G. I., “The spectrum of turbulence,” in [*Proc. R. Soc. London Ser. A*], (1938).
- [78] Antoniou, A. and Lu, W.-S., [*Practical Optimization: Algorithms and Engineering Applications*], Springer (2007).
- [79] Cottet, F., [*Traitement des signaux et acquisition de données*], Dunod (2002).
- [80] Assémat, F., Wilson, R., and Gendron, E., “Method for simulating infinitely long and non stationary phase screens with optimized memory storage,” *Optics Express* **14**, 988–999 (Feb. 2006).
- [81] Fried, D. L. and Clark, T., “Extruding kolmogorov-type phase screen ribbons,” *J. Opt. Soc. Am. A* **25**(2), 463–468 (2008).
- [82] Beghi, A., Cenedese, A., and Masiero, A., “Stochastic realization approach to the efficient simulation of phase screens,” *J. Opt. Soc. Am. A* **25**(2), 515–525 (2008).
- [83] Kulcsár, C., Raynaud, H.-F., Petit, C., Conan, J.-M., and de Lesegno, P. V., “Optimal control, observers and integrators in adaptive optics,” *Opt. Express* **14**(17), 7464–7476 (2006).
- [84] Petit, C., Conan, J.-M., Kulcsár, C., and Raynaud, H.-F., “Linear quadratic gaussian control for adaptive optics and multiconjugate adaptive optics: experimental and numerical analysis,” *J. Opt. Soc. Am. A* **26**(6), 1307–1325 (2009).

- [85] Clare, R. M., Ellerbroek, B. L., Herriot, G., and Véran, J.-P., “Adaptive optics sky coverage modeling for extremely large telescopes,” *Appl. Opt.* **45**, 8964–8978 (Dec 2006).
- [86] Rigaut, F. and Gendron, E., “Laser guide star in adaptive optics - the tilt determination problem,” *Astronomy and Astrophysics* **261**, 677–684 (1992).
- [87] Waddle, A., “Benchmarking GPU implementations of tomographic reconstruction algorithms for the TMTs NFIRAOS.” University of Victoria Dept. of CS and NRC Herzberg, Internal report (2012).
- [88] Sivo, G. et al., “First laboratory validation of LQG control with the CANARY MOAO pathfinder,” *Proc. SPIE* **8447**, 2Y (2012).
- [89] Conan, R., “Object-oriented Matlab adaptive optics,” *Proc. SPIE* **9148**, 249 (2014).
- [90] Tokovinin, A., “From differential image motion to seeing,” *PASP* **114** (2002).
- [91] Els, S. G., Travouillon, T., et al., “Thirty Meter Telescope Site Testing VI: Turbulence Profiles,” *PASP* **121**(879) (2009).
- [92] Miyashita, A., Takato, N., et al., “Statistics of the weather data, environment data, and the seeing of the Subaru Telescope,” *Proc. SPIE* **5489**, 207 (2004).
- [93] Costille, A. and Fusco, T., “Impact of the Cn2 description on WFAO performance,” *Proc. AO for ELTs* **2** (2011).
- [94] Gendron, E., Martin, O., Gratadour, D., Vidal, F., and Rousset, G., “Robustness of tomographic reconstructors versus real atmospheric profiles in the ELT perspective,” *Proc. SPIE* **9148**, 179 (2014).
- [95] Brenner, J. and Cummings, L., “The hadamard’s maximum determinant problem,” *Am. Math. Monthly* **79**, 626–630 (1972).
- [96] Pham, L., *Optimizing the Optical Calibration Performance of a Multi-Object Adaptive Optics Instrument*, Master’s thesis, University of Victoria (2013).

## Appendix A

# Performance Modeling for the RAVEN Multi-Object Adaptive Optics Demonstrator

## Performance Modeling for the RAVEN Multi-Object Adaptive Optics Demonstrator

DAVID R. ANDERSEN,<sup>1</sup> KATE J. JACKSON,<sup>2</sup> CÉLIA BLAIN,<sup>2</sup> COLIN BRADLEY,<sup>2</sup> CARLOS CORREIA,<sup>1</sup> MEGURU ITO,<sup>2</sup>  
OLIVIER LARDIÈRE,<sup>2</sup> AND JEAN-PIERRE VÉRAN<sup>1</sup>

*Received 2011 September 15; accepted 2012 March 20; published 2012 April 25*

**ABSTRACT.** RAVEN will be a Multi-Object Adaptive Optics (MOAO) technology and science demonstrator on the Subaru telescope. The baseline design calls for three natural guide star (NGS) wavefront sensors (WFS) and two science pickoff arms that will patrol a  $\sim 2'$  diameter field of regard (FOR). Sky coverage is an important consideration, because RAVEN is both a technical and science demonstrator. Early-stage simulation of RAVEN's performance is critical in establishing that the key science requirement can be met. That is, 30% of the energy of an unresolved point-spread function (PSF) be ensquared within a 140 mas slit using existing WFS camera and deformable mirror (DM) technology. The system was simulated with two independent modeling tools, MAOS and OOMAO, which were in excellent agreement. It was established that RAVEN will be an order  $10 \times 10$  adaptive optics (AO) system by examining the tradeoffs between performance, sky coverage, and WFS field of view. The 30% ensquared-energy (EE) requirement will be met with three NGSs and will exceed 40% if the Subaru Laser Guide Star (LGS) is used on-axis (assuming median image quality). This is also true for NGSs as faint as  $m_R = 14.5$ .

*Online material:* color figures

### 1. INTRODUCTION

The dawn of Extremely Large Telescopes (ELTs) is upon us. The Thirty Meter Telescope (TMT; Nelson 2008), the European-ELT (E-ELT; Kissler-Patig 2010), and the Giant Magellan Telescope (GMT; Shectman & Johns 2010) are all approaching their construction phase. Near-infrared spectrographs with 20 or more deployable integral field units over a 5 to 10' FOR, assisted by MOAO, are highly desirable potential instruments on ELTs, because they can be used to address major areas in their top-level science cases. These MOAO integral field spectrographs (IFSs) are ideally suited for studying the evolution of galaxies from first light to the era of peak star formation. However, use of such an instrument will not be limited to extragalactic astronomers; any astronomer seeking multi-object spectroscopy that takes advantage of the “ $D^4$ ” sensitivity gain provided by AO will consider a MOAO IFS to be a work-horse instrument. Infrared Multi-Object Spectrograph (IRMOS; Gavel et al. 2006; Andersen et al. 2006) and Extremely Large Telescope Adaptive Optics for Galaxy Evolution (EAGLE; Cuby et al. 2010) instrument are two examples of MOAO + IFS instrument concepts for, respectively, the TMT and the E-ELT.

To achieve correction over a large FOR, a MOAO system must overcome an effect known as “anisoplanatism.” For a given telescope pointing, the light from a distant source is perturbed by the turbulence in a cylinder (with a diameter the size of the telescope primary mirror). Light from a nearby source will pass through an overlapping, but nonidentical, cylinder of turbulence on its way to the telescope. In a classical AO system, a single WFS will pick off light from a single, relatively bright, point source, and a DM will be commanded to take the appropriate shape to null-out the wavefront error induced by the turbulence along a single line of sight (within a single cylinder). The AO correction for a different source will not be as good, because it will be viewed through a slightly different cylinder of turbulence. Definitions vary slightly, but the isoplanatic angle,  $\theta_0$ , can be thought of as the angular distance from the guide star at which the Strehl ratio drops significantly. The quantity  $\theta_0 \propto \lambda^{6/5}$  and is typically 10" in the  $H$  band (for a corrected field of view [FOV] of  $\sim 20''$ ).

There are two approaches for enlarging the isoplanatic angle. One approach is to place multiple DMs in series, each conjugate to a different atmospheric altitude. This multiconjugate AO (MCAO; Johnstron & Welsh 1991; Ragazzoni 1999; Flicker et al. 2000) approach can be used to enlarge the FOV to sizes of an arcminute or two, but the performance will ultimately still be limited by generalized anisoplanatism (Rigaut et al. 2000). The FOV can be further enlarged by adding even more DMs in series, to remove the turbulence generated at even more

<sup>1</sup>National Research Council, Herzberg Institute of Astrophysics, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada.

<sup>2</sup>Adaptive Optics Laboratory, Department of Mechanical Engineering, University of Victoria, Victoria, BC V8P 5C2; david.andersen@nrc-cnrc.gc.ca.

atmospheric heights, but the complexity of the MCAO system rises (and the throughput falls) with each additional DM relay.

MOAO is a parallel approach that promises to increase the field over which AO corrections can be applied to 5 or even 10' (Hammer et al. 2002). MOAO systems use the fact that there are only a limited number of interesting targets in a given FOR, and astronomers will be happy if AO corrections are made only in those directions. If a sufficiently accurate measurement is made of the turbulent volume over a telescope, one can place a probe with an embedded DM anywhere in the FOR and make the optimal turbulence correction for that position. Making a measurement of the turbulent volume requires information from multiple WFSs locked on multiple guide stars that probe different lines of sight through the atmosphere. Once the information from these multiple WFSs is combined into a single tomographic model of the turbulence (Tokovinin et al. 2001), it is straightforward to imagine multiple science pickoffs in parallel, each incorporating its own DM, feeding multiple IFSs. Falcon for VLT was the first proposed MOAO IFS (Hammer et al. 2002; Puech et al. 2006), and it has served as a model for the more recent IRMOS and EAGLE studies for ELTs.

Many of the challenges involved in designing a MOAO system, such as the use of tomography (Ragazzoni et al. 1999; Costille et al. 2010; Ammons et al. 2010), microelectromechanical system (MEMS) mirrors (Morzinski et al. 2010), and woofer-tweeter control (Jackson et al. 2010), have all been demonstrated to work in different laboratory settings and are included in advanced instrument concepts. Open-loop (OL) control is perhaps the greatest risk to MOAO, however, partly because it is the biggest unknown. In an AO system with OL control, the WFSs do not sense the correction applied by the DM. Instead, the WFSs sense the full turbulent phase of the atmosphere and the DMs are commanded to take the appropriate shape without benefit of any feedback. While OL control is not a new idea (Primmerman et al. [1991] used so-called go-to adaptive optics to make corrections and take science images immediately following pulses from a laser guide star with a low duty cycle), interest in implementing open-loop control on-sky has been reinvigorated in the past few years, as we shall see in the next section. After all, OL control introduces unique requirements on an AO system: the WFS needs to have a high dynamic range; effects of DM hysteresis and nonlinearity need to be mitigated; and, finally, alignment and calibration become more challenging.

### 1.1. MOAO Demonstrators

While the risks associated with MOAO IFSs have kept proposed VLT and ELT instruments on the drawing board, the scientific promise is so great that multiple on-sky demonstrators have been developed. The Visible Light Laser Guidestar Experiments (ViLLaGEs; Gavel et al. 2008; Ammons et al. 2008) is a MEMS DM-based AO testbed on the Nickel 1 m telescope at Lick Observatory. ViLLaGEs carried out on-axis experiments in both

closed and open loop with NGSs and LGSs. It was the first on-sky experiment to successfully demonstrate open-loop control. ViLLaGEs is a test bed that is being employed to develop the Keck Next Generation Adaptive Optics (NGAO; Wizinowich et al. 2010; Ammons et al. 2010) instrument, which is a tomographic, high-order, open-loop AO system.

The Victoria Open Loop Testbed (VOLT; Andersen et al. 2009) was an experiment aimed at distilling the problems of open-loop control into a simple experiment. VOLT demonstrated open-loop control in the laboratory and on-sky at the Dominion Astrophysical Observatory 1.2 m telescope using a simple on-axis NGS system (Andersen et al. 2008). Both the VOLT and ViLLaGEs open-loop AO demonstrators performed below expectations at low temporal frequencies, which seems to indicate that small misalignments in open-loop AO systems may ultimately limit their performance. These experiences have led to a second generation of MOAO demonstrators that emphasize both calibration and alignment techniques.

CANARY is a MOAO demonstrator at the William Herschel Telescope (Vidal et al. 2010; Gendron et al. 2010; Morris et al. 2010) that is considered a pathfinder for EAGLE on the E-ELT. The goals of the CANARY project are to perform NGS-based (and, subsequently, LGS-based tomographic wavefront sensing), perform open-loop AO correction on-sky, and develop calibration and alignment techniques. This experiment saw first light in the fall of 2010 and achieved a MOAO Strehl ratio of 26% in the *H* band (Gendron et al. 2011). CANARY will ramp up to a full MOAO test bed with multiple LGSs by 2013. While the performance of CANARY at low temporal frequencies was improved, it still suffered in comparison with the performance in closed loop (E. Gendron 2011, private communication).

### 1.2. The RAVEN MOAO Demonstrator

RAVEN will be the first MOAO instrument on an 8 m class telescope feeding an AO-optimized science instrument, the Subaru Infrared Camera and Spectrograph (IRCS; Tokunaga et al. 1998).<sup>3</sup> RAVEN has many of the same technical aims as CANARY, but also has some significant differences. Figure 1 shows a functional block diagram for RAVEN.

RAVEN consists of nine main subsystems:

1. The deployable calibration unit (CU) is a telescope simulator and a turbulence generator. It contains an array of off-axis NGS sources and one on-axis LGS source. Light from the CU will feed the three OL WFSs, the LGS WFS, and two science arms. The three functions of the CU are to (1) help align other RAVEN subsystems, (2) calibrate the AO system (generate interaction matrices and measure field-dependent non-common-path aberrations), and (3) test the MOAO system by

<sup>3</sup> See <http://www.naoj.org/Observing/Instruments/IRCS/>.

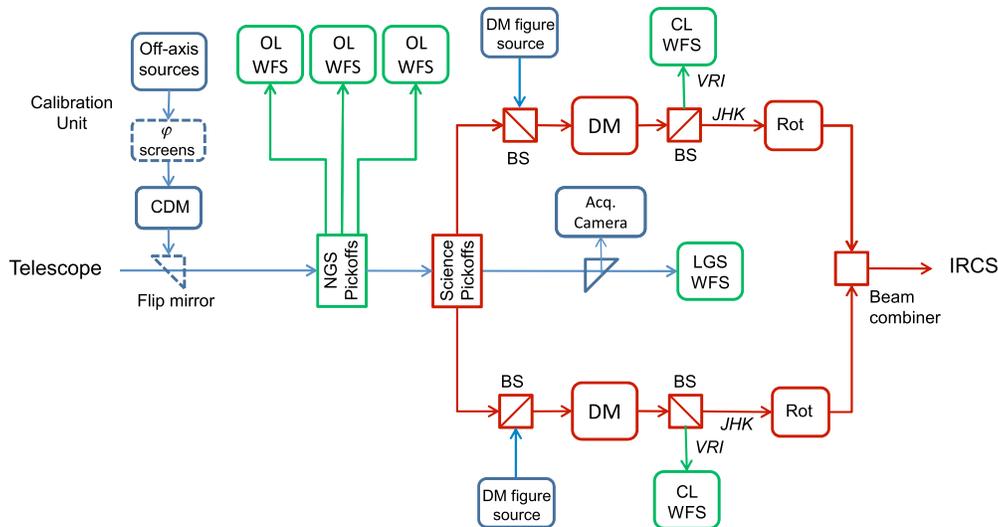


Fig. 1.—Functional optical block diagram of RAVEN. Dashed blocks are deployable. RAVEN consists of nine main subsystems: the deployable calibration unit, the open-loop NGS WFSs, the science pickoffs, the science relays, the closed-loop NGS truth/figure WFSs, the beam combiner, the LGS WFS, and the acquisition camera. The real-time computer is not shown. See the electronic edition of the *PASP* for a color version of this figure.

including three phase screens (including a ground-conjugate DM inside the CU).

2. The three NGS OL WFSs are mounted on  $x$ - $y$  translating stages to prevent the pupil from rotating on the WFS lenslet array with respect to the DMs.

3. RAVEN includes an on-axis LGS WFS, which will be fed by the Subaru sodium beacon in order to improve AO correction and/or the sky coverage (as discussed in § 3.5.2).

4. The science pickoff design consists of a mirror mounted on a  $r$ - $\theta$  arm, followed by a trombone mirror that keeps the path length constant.

5. The science relay for each arm contains a DM (which we expect to be a custom ALPAO DM with  $11 \times 11$  actuators with a 25 mm aperture).

6. A figure source and closed-loop (CL) WFS share the science relay optical path and can be used to either (1) measure the shape of the DM using the figure source, (2) use the CL WFS as a truth WFS to help calibrate RAVEN or measure the MOAO performance, or (3) use the CL WFS as a classical AO system that uses the science target as the NGS.

7. After the science relay, light from both arms of the system is combined so that the common beam shares an identical exit pupil and provides two adjacent 4" science fields to the single IRCS slit. The beam combiner also contains two  $K$  mirrors, which can rotate the images of the science targets so that extended objects can be properly aligned onto the slits.

8. An acquisition camera can be used to determine the telescope pointing and ensure that shadows of the probe arm fall over the NGSs and science targets.

9. Finally, pixels from the WFS detectors will be read by the RAVEN real-time computer (RTC) and transformed into a

tomographic model of the atmosphere above the observatory. This tomographic model will be sampled in directions defined by the position of the science probes in the patrol field, and DM commands will be generated and applied.

The science gain achievable by RAVEN, in comparison with classical AO systems such as Subaru's AO188 (Minowa et al. 2010), will be modest, because RAVEN will only have two science channels. Nevertheless, the 8 m aperture of the Subaru telescope enables science that is not achievable on smaller telescopes, and RAVEN will be capable of delivering high ensquared energy into the IRCS slit. The combined technical and scientific aspects of MOAO that RAVEN will demonstrate are meant to excite the astronomical community and build support for future facility-class MOAO instruments with much larger multiplex advantages for either 8 m class telescopes or ELTs.

### 1.3. RAVEN Performance Modeling

MOAO has the potential to deliver near-diffraction-limited images to multiple small patches spread across a large FOR. One challenge of an MOAO system is that it is highly distributed. For RAVEN, light from three or four guide stars will be sensed by open-loop WFSs and a tomographic model of the atmosphere generated by the RTC. The RTC will then produce DM commands specific to the direction of the science pickoff arms. All of these actions are performed using OL control. Accurate knowledge of the science probe's placement in the focal plane and the relative alignment of the DM and WFS in the pupil plane is required.

A broad swath of parameter space has been explored in order to determine if RAVEN can realistically meet the proposed performance requirements and deliver useful MOAO-corrected images to the Subaru IRCS spectrograph. As RAVEN was conceived to be a science-capable, NGS-only MOAO system, in addition to a technical demonstration, the AO architecture was designed such that it will deliver the desired performance when even faint guide stars are used. The addition of the single on-axis Subaru LGS to the NGS constellation improves performance and sky coverage, but does not eliminate the need for good performance with faint NGSs.

In this article, we assess the performance of the RAVEN MOAO system. In § 2, we describe MAOS (Multithreaded Adaptive Optics Simulator) and OOMAO (Object-Oriented MATLAB Adaptive Optics), the two simulation tools used in this study. In § 3, we describe the baseline system performance of RAVEN. This involves a description of the performance metrics, the derived atmospheric profiles, and the RAVEN ensquared-energy error budget. In § 4, we describe the trade studies that led us to our baseline design and describe RAVEN performance when used with different model atmospheres or at different zenith angles. Finally, we summarize our findings and present a road map of more detailed simulations that will be undertaken as we continue the development of RAVEN.

## 2. SIMULATION TOOLS

Our analysis was undertaken using two simulation platforms: MAOS and OOMAO.

### 2.1. MAOS

MAOS is a new *C* implementation of the tomographic AO simulator LAOS, which was written in MATLAB. This tool was created to efficiently run multithreaded simulations of large AO systems. MAOS is an ideal modeling tool for RAVEN, because it fully implements zonal tomography and can be configured for completely open-loop MOAO operations.<sup>4</sup>

The atmospheric turbulence in the model is composed of multiple phase screens that are located at different altitudes. These independent phase screens can either translate across the pupil, assuming a frozen flow for a given wind speed and direction, or MAOS can be run in a temporally uncorrelated mode in which each step of the simulation is temporally independent of the last (i.e., a simulated atmospheric phase is sensed by the WFSs and corrected by the DM in a single step). Running RAVEN simulations using this mode is advantageous, because the temporal errors are small (when run at a 500 Hz sampling frequency) compared with the tomographic errors, and the time-averaged PSF is more uniform. The resulting aberrations, due to

the simulated atmosphere, are sensed by multiple NGS (and LGS) Shack-Hartmann WFSs. MAOS can simulate the full physical optics WFS model that uses input pixel characteristics. Centroiding is accomplished using an optimal matched-filter (MF) estimation algorithm (Gilles & Ellerbroek 2006). The LGS WFS includes LGS elongation and cone effect. The tomographic wavefront reconstruction estimates the turbulence at several different heights from the open-loop gradients measured by the RAVEN WFSs, using a computationally efficient implementation of a minimum-variance reconstruction algorithm (Ellerbroek 2002). The reconstructed turbulent atmosphere is then projected, in MOAO mode, onto a DM corresponding to a given field direction. MAOS provides the user with numerous performance metrics, including rms wavefront error (both total and tip/tilt removed), Strehl ratio, and PSFs (defined at a given wavelength) for numerous field (DM) locations. Because MAOS is a highly optimized and proven tomographic Monte Carlo simulation platform, it was used for most simulations in this article.

### 2.2. OOMAO

The OOMAO modeling library is a set of MATLAB classes developed for the purpose of facilitating a clear and accessible end-to-end model of the RAVEN system. Objects from the different classes are assembled to perform numerical modeling of an AO system. OOMAO can be seen as an extension of the MATLAB language; overloaded MATLAB operators are used to propagate the wavefront through the system and to update the status of each object.<sup>5</sup>

Asterisms can be defined using the source class, with any number of guide stars in constellations specified by altitude-azimuth coordinates. The source class has a very important role in the OOMAO library, as it is the link between other classes. A source object carries a wavefront, both amplitude and phase, through the different objects representing the atmosphere, the telescope, the wavefront sensor, etc. Both NGS and LGS asterisms can be simulated. Currently all guide stars in the same asterism are defined with the same wavelength. Science source objects can be defined individually, assigned their own magnitude and wavelength, and placed at any point in the field of regard.

A modal tomography algorithm is implemented to reconstruct the phase, along with a thresholded center of gravity (CoG) to measure the spot positions on the WFSs. Using these methods, an end-to-end open-loop model of RAVEN has been developed with movable science objects in an adjustable asterism.

<sup>4</sup>MAOS is available from <https://github.com/lianqi/maos/>. MAOS is written by Lianqi Wang and developed by Lianqi Wang, Luc Gilles, and Brent Ellerbroek of the TMT AO group.

<sup>5</sup>OOMAO is available from <https://github.com/rconan/OOMAO>. OOMAO was originally developed by Rodolphe Conan. Peter Hampton, Kate Jackson, and Olivier Lardière provided additional contributions.

Various classes have been modified to include functionalities that will facilitate the simulation of error sources such as misalignment. The DM class includes the ability to misalign the DM with respect to the optical path and other system components by specifying five parameters: horizontal and vertical displacements, rotation, tip, and tilt. Sensor noise has been included via the detector subclass. Parameters include readout gain, thermal dark signal, excess noise factor, charge capacity, and clock-induced charge, which allows for easy modeling of electron-multiplying CCDs (EMCCDs). The frame rate and exposure time are decoupled (i.e., the exposure time can be shorter than the time between frames). An aberration object uses Zernike modes to create a static or quasi-static aberration at a selected point in the optical path. This object can also be defined with a diameter that is much larger than that of the optical path so that larger optics with static aberration can be modeled easily. The `RotateDisplace` object shifts a specified image by a given number of pixels and then rotates by a given number of radians and is used within the aberration object to shift the large optical aberration across the smaller optical path. In this way, aberrations can be introduced at various points in the simulated optical path that are representative of potential errors in the real optical system. As shall be shown in the next section, OOMAO can reproduce the simulation results. This verification is important, as it builds confidence in the results from both simulation tools. It is especially important for us to trust the OOMAO simulations, because we intend to control real RAVEN hardware using OOMAO in the early stages of testing the instrument (before the RAVEN RTC is completed).

### 3. BASELINE SYSTEM PERFORMANCE OF RAVEN

As will be shown in the following sections, tomographic errors are the dominant factor limiting the performance of RAVEN. As a result, the performance will be highly dependent on the total amount of turbulence (and the distribution of turbulence as a function of altitude) and on the asterism of NGSs used to sense the turbulence. In this section, the metrics used to evaluate the performance are described, and the atmospheric profiles derived to use in simulation are outlined. The ensquared-energy budget of RAVEN and an estimate of the system's limiting magnitude are also presented.

#### 3.1. Performance Metrics

RAVEN will feed the Subaru IRCS infrared imager and slit spectrograph. Since the majority of RAVEN science will be performed using the spectrograph, ensquared energy within the slit will be the most important performance metric. The IRCS echelle slit width is 140 mas wide, so ensquared energy (EE) within 140 mas at a wavelength of  $1.65 \mu\text{m}$  ( $H$  band) was used as the primary performance metric. EE was evaluated primarily in the  $H$  band, because RAVEN will have a slightly higher thermal background in the  $K$  band relative to AO188, due to the

increase in the number of optical elements, and the performance in the  $H$  band will obviously be better than in the  $J$  band, due to the longer wavelength. The requirement for ensquared energy in the  $H$  band was set to 30% in order to match the requirement on ensquared energy from AO188.<sup>6</sup> By meeting this requirement, the multiplex advantage of RAVEN will be very nearly double that of AO188 (assuming that the science channel throughput of RAVEN continues to be greater than 40%).

Since 140 mas is significantly wider than the 42 mas diffraction-limited spot at  $1.65 \mu\text{m}$ , RAVEN performance is most dependent on high spatial order wavefront errors (WFEs) and is relatively immune to modest errors at low spatial frequencies, including tip/tilt and focus. Therefore, another useful metric to evaluate in the simulations is the tip/tilt-removed WFE (other low-order aberrations, such as focus, could also be excluded, in principle). The Strehl ratio is also calculated (again measured at the  $H$  band), as is the total WFE, as these will be the quantities of interest when RAVEN is used with IRCS in imaging mode.

We have focused on setting the basic system parameters, such as system order, field of regard, and the limiting magnitude. To understand how performance varies with these and other parameters, we defined an asterism with three NGSs on a ring of  $45''$  radius and then evaluated the performance at multiple field points within that circle (in some simulations, a LGS WFS was included at the field center). The average performance is defined over points out to  $30''$  from the field center (excluding the field center when a LGS was used in the simulation).

#### 3.2. Atmospheric Profiles

We derived atmospheric profiles for our RAVEN simulations by combining image-quality measurements from the Subaru Observatory and differential image motion monitor (DIMM) and multiaperture scintillation sensor (MASS) turbulence profiles measured by TMT at Mauna Kea. We used a realistic seven-layer profile generated from the TMT MASS/DIMM site survey MK 13N data (Els et al. 2009). To assemble representative profiles corresponding to the quartiles of this residual WFE, more than 10,000 individual MASS/DIMM profiles were sorted by uncorrectable residual WFEs (fitting-plus-lag error), and 10% of the profiles clustered around the quartiles were averaged (M. Schoeck 2011, private communication). While uncorrectable residual WFE is not equivalent to image quality (IQ), it is an acceptable surrogate. These three seven-layer turbulence profiles (starting 60 m above the MK 13N site), are given in Table 1.

We complemented these MASS/DIMM profile measurements with measures of Subaru IQ taken between 2000 and 2004 (Miyashita et al. 2004). Assuming an infinite outer scale,

<sup>6</sup>See the figures at <http://www.naoj.org/Observing/Instruments/AO/performance.html>.

TABLE 1  
SUBARU ATMOSPHERIC PROFILES USED FOR RAVEN  
PERFORMANCE MODELING

$h$ (km)	$J^2$ 25%	$\int C_n^2 dh$ 50%	$(m^{1/3})$ 75%
0	$4.380 \times 10^{-14}$	$9.419 \times 10^{-14}$	$9.991 \times 10^{-14}$
0.06	$7.345 \times 10^{-14}$	$1.0318 \times 10^{-13}$	$1.5225 \times 10^{-13}$
0.5	$1.407 \times 10^{-14}$	$3.190 \times 10^{-14}$	$6.990 \times 10^{-14}$
1	$4.882 \times 10^{-15}$	$1.077 \times 10^{-14}$	$2.919 \times 10^{-14}$
2	$3.956 \times 10^{-14}$	$1.233 \times 10^{-14}$	$3.249 \times 10^{-14}$
4	$1.744 \times 10^{-14}$	$2.879 \times 10^{-14}$	$4.212 \times 10^{-14}$
8	$1.118 \times 10^{-14}$	$2.264 \times 10^{-14}$	$4.525 \times 10^{-14}$
16	$2.612 \times 10^{-14}$	$2.734 \times 10^{-14}$	$3.538 \times 10^{-14}$
$r_0$ (500 nm)	19.4 cm	15.6 cm	12.1 cm
FWHM	0.53"	0.66"	0.85"

we determined the Fried parameter,  $r_0$ , for each quartile of IQ (Table 1).<sup>7</sup>

The image-quality-derived Fried parameter values are smaller than the TMT site-testing  $r_0$  values derived from

$$r_0 = (16.7/\lambda^2 \cos^{-1} \gamma \sum_i J_i^2)^{-3/5}, \quad (1)$$

where  $J_i^2 = \int C_n^2(h) dh$  over the  $i$ th layer, and  $\gamma$  is the zenith angle. We interpret these differences as being due to the local ground layer at the Subaru telescope, wind shake, and dome seeing. While not all of these PSF-broadening terms will necessarily follow a von Kármán spectrum, assuming that all the differences in  $r_0$  are due to an additional ground-layer term that follows von Kármán is probably a conservative assumption. The first row, corresponding to 0 m, in Table 1 corresponds to this additional turbulence required to decrease the TMT site-testing-derived  $r_0$  values to be in line with the Subaru image-quality measurements.<sup>8</sup>

This additional dome/ground seeing component in these profiles skews the fraction of the turbulence below 1 km to be greater than 60%. If the dome/ground seeing component is smaller and the free atmospheric turbulence is stronger (i.e., the isoplanatic angle is smaller), there will be a significant impact on RAVEN performance.

### 3.3. Simulated Performance

We simulated the performance of RAVEN using the median Subaru profile defined above. The three-NGS WFSs were on a 45" radius ring and had 10 subapertures across the pupil. Section 4.3.1 shows the results of a system-order trade study that

<sup>7</sup>The true value of  $r_0$  is probably overestimated by 10 to 20% by assuming an infinite outer scale (Martínez et al. 2010).

<sup>8</sup>In practice, the 0 and 60 m layers were combined into a single ground-layer profile in the simulations.

concludes that little improvement in performance can be gained for higher-order systems. A study of asterism diameter in § 4.1 also shows that there is limited improvement for guide star separations greater than 2' diameter, as the footprints of the NGSs in the metapupils separate at relatively low altitudes, leaving much of the atmospheric turbulence unsensed and uncorrectable. The selected configuration parameters for the baseline design are given in Table 2.

Initial simulations of this baseline system show the best performance possible for RAVEN, since WFS noise and other implementation errors are not included (Fig. 2). The figure only shows the performance at points far from the NGSs. At the location of the NGSs, the WFE is  $\sim 180$  nm rms, which is consistent with the fitting error for a  $11 \times 11$  order DM with a  $r_0$  of 15.6 cm (145 nm). The fact that the wavefront error over most of the field is substantially higher than this ( $\sim 270$  nm) suggests that the tomographic error is the dominant error source for RAVEN. It is due, in large part, to the incomplete overlap of the guide star footprints in the metapupils at higher altitudes, and this source of tomographic error will therefore decrease for larger-diameter telescopes (Tokovinin et al. 2001). Because the tomographic error was large, it was not essential to minimize the fitting error. Therefore, we adopted a relatively low order WFS (the 0.8 m  $d_0$  value is quite large for most AO systems),

TABLE 2  
RAVEN BASELINE CONFIGURATION PARAMETERS

Parameters	Values
Telescope	
Diameter	8 m
Central obscuration	2 m
Atmosphere	
$r_0$ (500 nm)	15.6 cm
$L_0$	30 m
Zenith angle	0°
Profile	Subaru 50%
wind <sub>GL</sub>	5.6 m s <sup>-1</sup>
wind <sub>top</sub>	19.1 m s <sup>-1</sup> (at 8 km)
wind <sub>dir</sub>	random
Sampling	1/64 m
Wavefront Sensor	
$N_{\text{NGS}}$	3
NGS radii	45"
Order	$10 \times 10$
$\theta_{\text{pix}}$	0.4"
$N_{\text{pix}}$	15
$f_{\text{sample}}$	500 Hz
$\lambda_{\text{WFS}}$	0.7 $\mu\text{m}$
DM	
Order	$11 \times 11$
Stroke	Infinite
Influence functions	Bicubic spline
Evaluation	
$N_{\text{points}}$	49
$\lambda_{\text{evl}}$	1.65 $\mu\text{m}$
Sampling <sub>PSF</sub>	$\lambda/4/D$

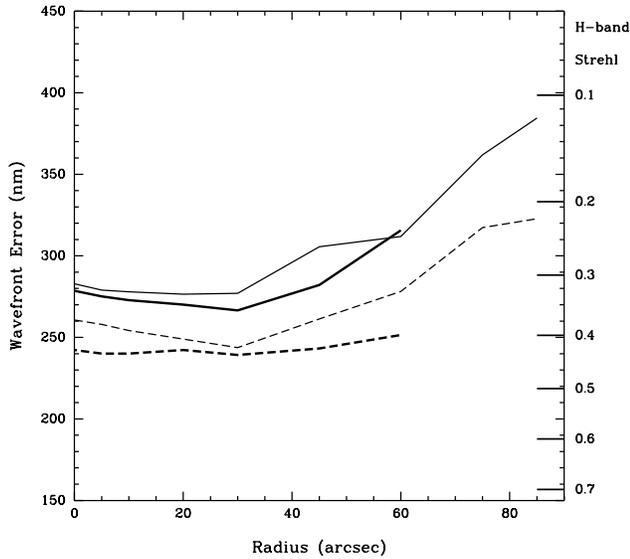


FIG. 2.—RAVEN performance for the baseline NGS-only configuration. Mean wavefront errors (all modes: *thin solid line*; tip/tilt removed: *thin dashed line*) vs. radius for field locations sampling half the focal plane for three NGSs on a 45" radius ring. Mean fractional *H*-band EE (within 140 mas; *heavy dashed line*) and Strehl ratios (*heavy solid line*) measured from the PSFs are shown with the scale on the right (scaled to WFEs for comparison purposes by employing the Maréchal approximation,  $SR \approx \exp(-\omega^2)$ , where  $\omega$  is the rms wavefront error in radians.) Only field locations not directly adjacent to the NGSs were considered.

which will allow us to maintain good corrections even when using a relatively faint NGS. This system should exceed the RAVEN 30% EE requirement by delivering  $\sim 40\%$  EE within 140 mas, while delivering  $\sim 30\%$  Strehl ratios over most positions in the field (in the absence of implementation errors discussed in § 3.4).

### 3.3.1. Comparison of MAOS and OOMAO Results

The predicted RAVEN performances from the MAOS and OOMAO simulations of the baseline system (Table 2) are in excellent agreement. Both simulation tools predicted a mean Strehl ratio of 30% for the points within 30" of the field center. The two simulations also predicted an identical 43% ensquared energy in this area. Only the mean WFE differed slightly; MAOS predicted a slightly lower mean WFE of 290 nm, while OOMAO predicted 300 nm rms of WFE. The minimum WFE, reached at the NGS radius, is 220 nm rms. This is higher than the 180 nm rms best-case WFE obtained with MAOS. It is suspected that either the thresholded CoG used by OOMAO gives slightly poorer performance than the MF used in MAOS under these conditions, or the finite number of radial orders (8 in this case) used in the OOMAO tomographic reconstructor limits the simulated performance. Overall, however, the excellent agree-

ment between these two independent AO simulation tools provides us with confidence in our results and signals that both tools can be used interchangeably in our RAVEN simulations.

### 3.3.2. RAVEN Performance Gain from the Subaru LGS

RAVEN benefits greatly if it makes use of the Subaru LGS facility. The NGS WFSs can be moved farther out while still covering a majority of the metapupil at the top of the turbulent atmosphere (the NGS could be moved out onto a  $\sim 3'$  diameter ring while maintaining some overlap with the LGS metapupil at 16 km). This will improve sky coverage, because the area over which NGS can be chosen while still meeting the RAVEN performance requirements will be substantially larger. Additionally, RAVEN can work with one LGS and just two NGSs (over a 2' FOR). Again, this will markedly improve sky coverage if only two NGSs are required. The LGS is relatively bright; this can help compensate the AO performance of RAVEN when the other NGSs are faint, and superior performance can be achieved if three bright NGSs are found within a 2' ring. Median WFE within a 2' field will be  $\sim 230$  nm ( $\sim 190$  nm of tip/tilt-removed wavefront error; Fig. 3). This corresponds to a large increase in ensquared energy (more than 50%) and Strehl ratio (greater than 50% up to 20" from the LGS). These performance

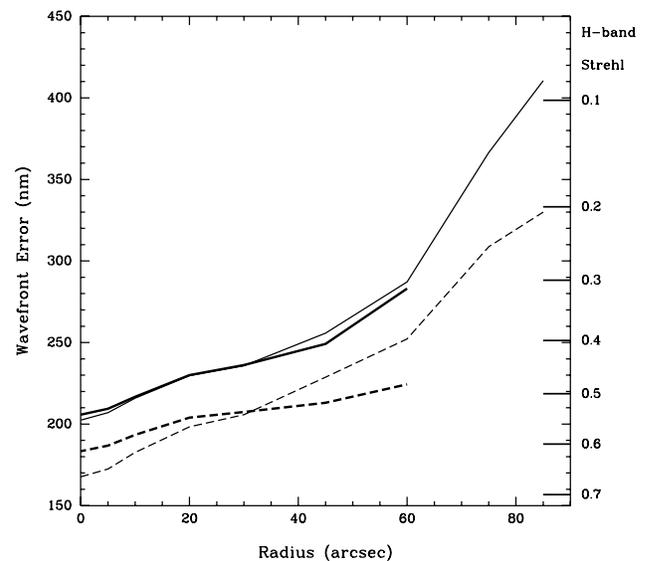


FIG. 3.—RAVEN performance for the baseline three NGSs plus on-axis LGS configuration. Mean wavefront errors (all modes: *thin solid line*; tip/tilt removed: *thin dashed line*) vs. radius for field points sampling half the focal plane for three NGSs on a 45" radius ring and an on-axis LGS. Mean fractional *H*-band ensquared energy (within 140 mas; *heavy dashed line*) and Strehl ratios (*heavy solid line*) measured from the PSFs are shown with the scale on the right (scaled to the WFEs by the Maréchal approximation). Only field locations not directly adjacent to the NGSs were considered.

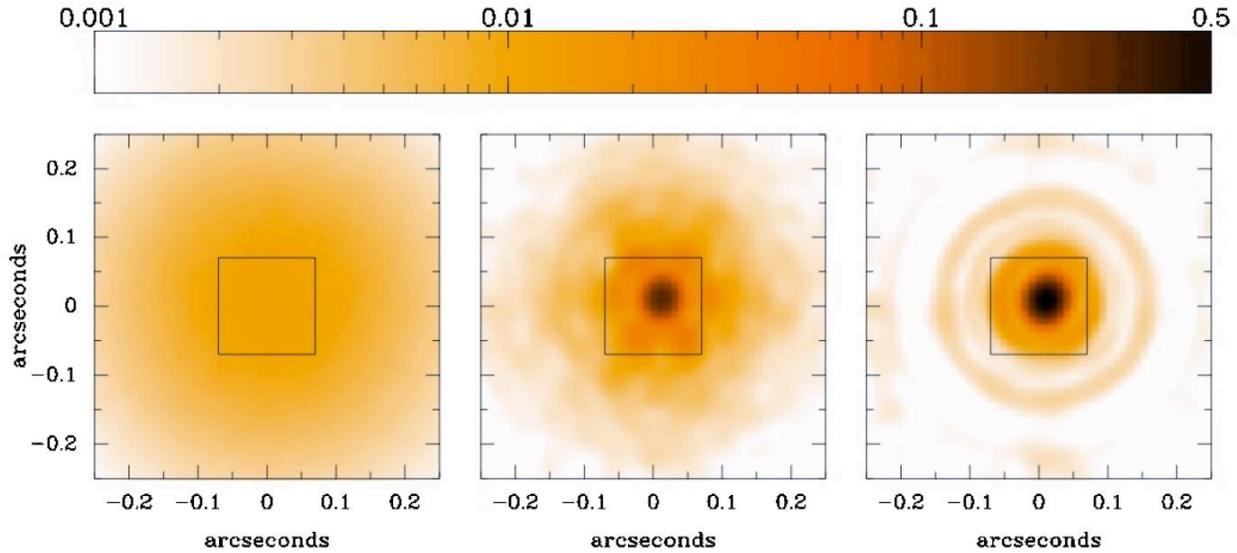


FIG. 4.—Simulated *H*-band PSFs in open-loop (*left*), RAVEN with three NGSs evaluated at the field center (*middle*), RAVEN with three NGSs and the LGS evaluated at the field center (LGS location; *right*). The box shows the 140 mas box in which ensquared energy is measured corresponding to the IRCS slit width; the simulated EEs are 10% with no AO, 40% with three NGSs, and 60% at the location of the LGS. All PSFs use the same stretch, with the scale corresponding to the peak flux of the perfect diffraction-limited PSF. See the electronic edition of the *PASP* for a color version of this figure.

predictions do not include the unmodelable implementation errors discussed in § 3.4.

A comparison of the open-loop (no AO correction) PSFs and the RAVEN-corrected PSFs (with and without the central LGS) is shown in Figure 4.

### 3.4. Error Budget

While the most important figure of merit for RAVEN is EE within a 140 mas slit, it is difficult to disentangle how potential sources of WFEs affect this metric. To give a sense of RAVEN performance, a WFE budget has been built in that it is easier for the expected contribution of different errors to be added together in quadrature. The high-order WFEs are of primary concern, because low-order errors (e.g., tip, tilt, and focus errors) will broaden the core of the PSF and not lead to significant losses in EE within a 140 mas box. The RAVEN WFE budget is listed in Table 3. The terms that are included in the simulations are described below, and then some additional implementation errors are listed; these include errors derived from laboratory tests of an ALPAO DM that is similar to the RAVEN science DMs.

**Tomographic Error.**—The dominant error term for RAVEN operating with just three NGSs will be the tomographic error. The median tomographic tip/tilt-removed WFE averaged over field points out to 40'' (not considering the center of the FOR when the LGS is considered) is 175 nm for RAVEN using just three NGSs on a 45'' radius ring and just 105 nm if the Subaru LGS is positioned at the center of the field. A certain fraction of this WFE is of relatively low order and will not substantially decrease the ensquared energy.

**DM Fitting Error.**—The  $11 \times 11$  RAVEN DM cannot be used to fit high-order modes and will therefore contribute a fitting-error term, equal to roughly

TABLE 3  
RAVEN TIP/TILT-REMOVED WAVEFRONT ERROR BUDGET

Term	Three NGSs	+LGS
Simulated WFE Terms (nm rms)		
Tip/tilt-removed tomography	175	105
DM fitting	155	145
WFS aliasing	103	112
WFS sampling	72	62
WFS noise ( $m = 12$ ; $F_s = 500$ Hz)	54	83
WFS noise ( $m = 14$ ; $F_s = 180$ Hz)	95	96
Simulated total ( $m = 12$ ; $F_s = 500$ Hz)	271	236
Simulated total ( $m = 14$ ; $F_s = 180$ Hz)	280	241
Simulated Ensquared Energy in 140 mas Slit		
Simulated EE ( $m = 12$ )	0.42	0.47
Simulated EE ( $m = 14$ )	0.40	0.46
Implementation WFE Terms (nm rms)		
Calibration +lag (−15 dB saturation; see text)	88	88
DM flattening	7	7
DM stability	27	27
DM repeatability	22	22
High-order optical errors	50	50
Implementation total WFE	107	107
Total WFE (nm rms)		
Total WFE ( $m = 12$ )	291	259
Total WFE ( $m = 14$ )	300	264
Total Ensquared Energy in 140 mas Slit		
Total EE ( $m = 12$ )	0.35	0.39
Total EE ( $m = 14$ )	0.33	0.38

$$\sigma_{\text{fit}}^2 \approx 0.25(d_0/r_0)^{5/3}, \quad (2)$$

where  $d_0$  is the interactuator spacing projected onto the primary mirror (0.8 m for the baseline system) and yields  $\sim 150$  nm of WFE. This makes fitting error the next greatest contribution to the high-order WFE budget.

**WFS Aliasing Error.**—The large size of the subapertures on the WFS will also contribute an aliasing error that arises from high spatial frequency disturbances that affect the WFS signal. From simulations, we find that this error can be characterized by  $\sigma_{\text{alias}}^2 \approx 0.1(d_0/r_0)^{5/3}$ , or  $\sim 100$  nm rms WFE, due to aliasing in the baseline system. This aliasing term will contribute to the scattering of light out of the PSF core and into the halo (and outside of the slit).

Taken together with the DM fitting error, the total generalized fitting error for RAVEN is large: approximately 180 nm rms. This is consistent with a generalized fitting error of roughly  $\sigma^2 = 0.35(d_0/r_0)^{5/3}$  and is the same as for classical AO systems. While the generalized fitting error is equal to the tomographic term for the NGS-only case, the generalized fitting error dominates the high-order error budget when the LGS is included as well. We maintain the large subapertures and relatively low order of the RAVEN MOAO system in order to achieve a higher sky coverage and WFS dynamic range, as discussed in §§ 3.5 and 4.3, respectively.

**WFS Sampling Error.**—This term refers to the WFE that arises from the undersampling of the WFS spots by the RAVEN WFSs. The WFSs need to be undersampled, due to the limited number of pixels available to cover an order  $10 \times 10$  system with a relatively large FOV. This term is dependent on the centroiding algorithm. We used the MF to determine this error. Other centroiding algorithms would have different WFS sampling and noise errors; we will choose a centroiding algorithm for RAVEN to minimize this error source.

**WFS Noise Error.**—The WFS noise was calculated from simulations with and without detector noise, photon noise, and sky background. This term obviously depends on guide star magnitude and the sampling rate. The noise errors quoted in Table 3 assume a 500 Hz sampling frequency for bright stars and a 180 Hz sampling frequency for fainter stars. This term can be reduced further if a slower sampling frequency is used at the expense of a larger lag error (which we include in the calibration-plus-lag error term of Table 3).

**Implementation Errors.**—There are a number of error sources that arise from sources that are not (yet) all simulated. These include the calibration-plus-lag error for an OL system; the DM flattening and go-to errors, as measured from the ALPAO DM 97 using a Zygo interferometer in the University of Victoria AO laboratory; and the high spatial frequency, uncorrectable errors on RAVEN optics. We made an educated guess that this last term will not exceed 50 nm rms WFE.

The top section of Table 3 contains the simulated error terms described above. The total errors for two different guide star

brightnesses (with and without the addition of a fourth, constant-brightness, LGS) are derived by adding the individual terms together in quadrature. The ensquared energies measured from the simulated PSFs are also tabulated. The bottom section of the table includes implementation errors not included in the end-to-end simulations. The total WFE is a quadrature sum of the simulated and implementation errors. It should be noted that the error terms and their values are consistent with those reported by the CANARY team (Gendron et al. 2010). The ensquared energy, accounting for implementation errors, is calculated by assuming that the loss of Strehl ratio due to the high-order implementation errors will remove the same amount of light from the 140 mas box.

In this RAVEN performance budget, it was found that the implementation errors will further reduce the ensquared energy by  $\sim 15\%$ . Most of this is due to the open-loop calibration and lag errors. We bundle these errors together by modeling the consequences of an open-loop rejection transfer function (RTF; the amount of turbulence that can be sensed and rejected as a function of frequency), which includes a constant wavefront error from misalignment or miscalibration. This constant error is manifested in the RTF as a plateau at low temporal frequencies; the amount of rejection over long timescales saturates. All open-loop AO systems to date have shown a RTF that saturates at low temporal frequencies. VOLT's RTF saturated at  $-15$  dB, while other systems performed somewhat better (see § 4.6 for more details). In Table 3, a conservative estimate of 88 nm rms was made by assuming that the RTF of RAVEN will saturate at  $-15$  dB. (This error only includes Zernike modes between 7 and 45. Errors on low-order Zernike modes will not substantially decrease the EE in 140 mas, and higher-order Zernike modes will not be fully corrected by RAVEN.) If a better calibration can be achieved, this saturation threshold may be reduced, perhaps to  $-25$  dB, and the corresponding WFE would drop from 88 nm to  $\sim 40$  nm. In this case, the implementation error would reduce the ensquared energy by just 10%. As long as no substantial implementation error remains unaccounted, the performance requirement of delivering 30% EE to the 140 mas spaxel will be met at zenith under median conditions. If the temporal error can be reduced through good calibration techniques, 40% EE can be achieved when using the Subaru LGS.

### 3.5. Limiting Magnitude of Asterism Guide Stars

The baseline RAVEN OL WFS detector is the EMCCD camera. These devices have the ability to operate with very high gain and low read noise. In this mode, the read noise can be made almost arbitrarily small; however, the background source plus background photon noise is effectively doubled. The signal-to-noise ratio (S/N) can be written as follows:

$$S/N = S_0 / \sqrt{(2(S_D + (S_0 + B)) + (N_a/G)^2)}, \quad (3)$$

where  $S_0$  is the source counts (all measured in electrons);  $B$  is the background;  $S_D$  is effectively the dark current;  $N_a/G$  is effectively the read noise divided by the gain, and this ratio is chosen to be 0.1; and the  $S_D$  for the Andor cameras is actually a combination of two effects: the thermal dark signal,

$$D_t = 3.3 \times 10^6 t^2 e^{-9080/T}, \quad (4)$$

which can effectively be ignored for short exposure times when the cameras are cooled.

### 3.5.1. NGS-only Case

Using the noise characteristics of this EMCCD device, simulations were run using the baseline configuration. First, simulations of RAVEN using only NGSs were performed. Each simulation evaluated the performance of RAVEN as we varied the NGS magnitudes (Table 4). The performance started to degrade significantly by  $m_R = 14.5$ , but almost all of that performance could be regained if RAVEN were run slower, at 180 Hz. The servo-lag error did not start impacting the simulated WFEs, but it was decided to require that the open loop run at a rate of at least 180 Hz (for these simulations), because MAOS does not account for an OL RTF that plateaus at low temporal frequencies (§ 4.6). If RAVEN can be run at 180 Hz without loss in performance, the limiting magnitude becomes 1 mag fainter. Figure 5 shows the simulated PSFs for faint-magnitude NGSs in comparison with a bright star. The bright core is still evident for the  $m_R = 14.5$  NGS asterism running at a sampling frequency of 180 Hz.

### 3.5.2. NGS with On-Axis LGS

RAVEN system performance improves in several ways with the addition of a central LGS. In addition to decreasing the tomographic error with the addition of another WFS (§ 3.3.2), RAVEN can use an asterism of three sources, including the LGS and just two NGSs. This will greatly enhance the sky coverage of RAVEN, as the NGS WFS probes can be moved 105'' away from the field center. Finally, the LGS can be used with three faint NGSs and still maintain significant EE. If it is assumed that the LGS is a  $m_R = 11$  beacon (Y. Hayano 2011, private communication), and the three NGSs are  $m_R = 15$ , then RAVEN can still deliver an average EE > 40% (before losses due to implementation errors) over a 1' diameter field (Fig. 6).

Sky coverage for RAVEN will be low, but the addition of the LGS to the NGS asterism will significantly improve the fraction of the sky that can be observed. As an example, consider a point with Galactic coordinates  $(b, l) = (30, 0)$ . Using the Besançon model of the Galaxy (Robin et al. 2003),<sup>9</sup> one finds that there are 750 stars per square degree with  $R < 14.5$  (1040 stars per

TABLE 4  
MEAN H-BAND PERFORMANCE OVER CENTRAL 30'' RADIUS  
AS FUNCTION OF NGS MAGNITUDE FOR NGS-ONLY CASE

$m_R$	$f_s = 500$ Hz		$f_s = 180$ Hz	
	EE <sub>140</sub>	Strehl	EE <sub>140</sub>	Strehl
Bright .....	0.43	0.30	...	...
10 .....	0.43	0.30	...	...
12 .....	0.42	0.27	...	...
14 .....	0.36	0.17	...	...
14.5 .....	0.33	0.12	0.40	0.18
15 .....	0.29	0.08	0.37	0.14
15.5 .....	...	...	0.34	0.11

square degree with  $R < 15$ ). The probability that there are three stars with  $R < 14.5$  in a 2' diameter FOR is just 3%. This does not even account for asterisms that are unsuitable; in some cases, the science targets will not be inscribed within the potential NGS asterism, and therefore the tomographic error will be too great. If one just requires two NGSs with  $R < 15$  within a 90'' diameter, with the third guide star provided by the LGS, the sky coverage increases to 10%. These asterisms are also more likely to be acceptable, as the observer will have greater flexibility in choosing the field orientation (and hence the on-sky coordinates of the LGS). If one can accept three NGSs with  $R < 15$  within a 3' diameter to be used in addition to the LGS, the potential sky coverage jumps to 34%. Again, a fraction of these potential asterisms may ultimately prove unacceptable for the given science targets, but at this Galactic latitude, the sky coverage will likely be 10 times higher with the LGS than without.

## 4. EXPLORING SIMULATION PARAMETER SPACE

We first focused on identifying and studying the major AO components: the WFSs and the DMs. Having established the basic AO architecture, the number of variables we explored was increased, and we included a closer examination of the possible consequences of open-loop temporal errors and different input model atmospheres.

### 4.1. Field of Regard with Respect to Asterism Geometry

Because RAVEN uses three NGS pickoffs that will patrol a FOR up to 3.5' in diameter, we simulated RAVEN performance using three NGSs on rings of different diameters. As expected, a reduction in the diameter of the asterism leads to improved performance in the area inscribed within the asterism, with a rapid falloff in performance outside. This is due to an increased overlap of the NGS metapupils at higher altitudes, which leads to a better tomographic estimate of the turbulence above the telescope. Because of the great flexibility of the RAVEN NGS pickoffs, the possible asterism geometries are practically limitless. These simple simulations show, however, that once the guide

<sup>9</sup>We generated synthetic catalogs of stars using <http://model.obs-besancon.fr/>.

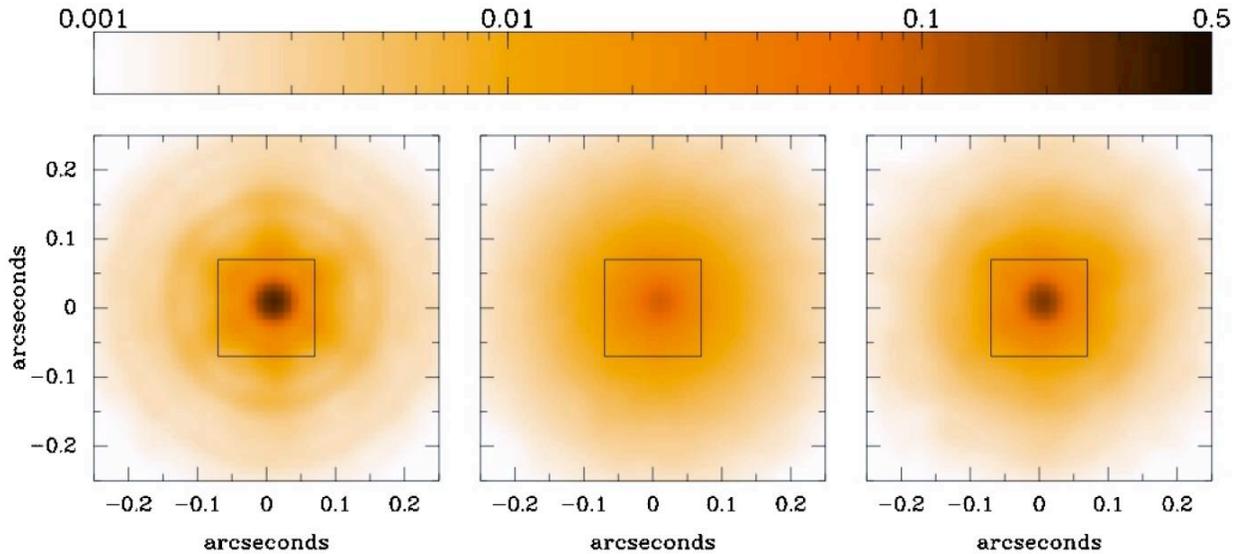


FIG. 5.—Comparison of field center PSFs for the three baseline NGS-only RAVEN simulations using guide stars with various magnitudes and sampling frequencies. *Left:*  $m_R = 10$  and  $f = 500$  Hz. *Middle:*  $m_R = 14.5$  and  $f = 500$  Hz. *Right:*  $m_R = 14.5$  and  $f = 180$  Hz. In the last of these cases, the core of the PSF is still present and a large fraction of the PSF energy remains within 140 mas (*boxed area*). See the electronic edition of the *PASP* for a color version of this figure.

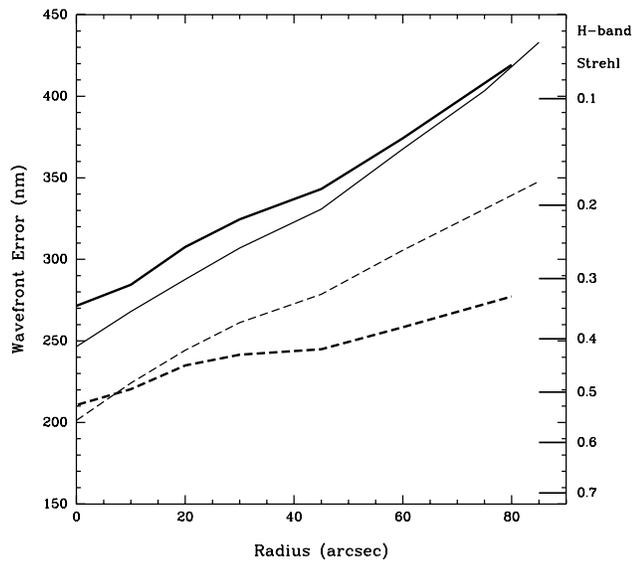


FIG. 6.—RAVEN performance for the three NGSs plus central LGS configuration when all three NGSs are faint ( $m_R = 15$ ) and the sampling frequency is reduced to 180 Hz. Wavefront errors (all modes: *thin solid line*; tip/tilt removed: *thin dashed line*) vs. radius for evaluation points covering half the FOR for three NGSs on a 45" radius ring. Mean fractional ensquared energy (within 140 mas; *thick dashed line*) and Strehl ratios (*thick solid line*) measured from the PSFs are shown with the scale on the right (scaled to the WFEs by the Maréchal approximation). Only field locations not directly adjacent to the NGSs were considered. The performance requirements of RAVEN should be met even when  $m_R = 15$  guide stars are used with the LGS.

stars are separated by  $2\lambda$ , the performance becomes very uniform within that diameter (Table 5).

#### 4.2. Enhanced Reconstruction and Control in Open Loop

RAVEN's performance and sky coverage could be improved by enhanced reconstruction and control techniques. Furthermore, it is believed that these methods will be critical to the success of future MOAO instruments on ELTs. Therefore, we advocate the use of RAVEN as a demonstrator in two areas of foremost importance for tomographic AO in general and MOAO in particular:

1. Operation in open loop makes it straightforward to use minimum pupil-integrated residual phase reconstructors (no temporal dynamics involved) in which the von Kármán spectrum is used as a priori knowledge for the volumetric estimation.

TABLE 5  
MEAN H-BAND PERFORMANCE OVER CENTRAL 15" RADIUS  
AS FUNCTION OF ASTERISM DIAMETER

Asterism diameter (arcsec)	Strehl ratio	Ensquared energy	WFE (nm)
30	0.41	0.47	256
45	0.41	0.46	256
60	0.39	0.46	260
90	0.30	0.43	302
120	0.27	0.42	311

In closed loop, however, to make use of this information, a technique called pseudo-open-loop control has been proposed, consisting of converting the closed-loop measurements back to open-loop by adding in the DM contribution (Gilles & Ellerbroek 2008). Clearly, this is now avoided. It thus makes RAVEN a perfect candidate to demonstrate the algorithms developed for closed-loop AO in a much more favorable terrain. Testing and use of high-performance iterative phase reconstructors as the conjugate gradient and block Gauss-Seidel have been pursued. We are currently assessing the optimal number of iterations and the real-time requirements for their implementation. Iterative algorithms embody an alternative to current vector-matrix-multiply reconstructors that use the generalized inverse of the interaction matrix between measurements and voltages. In addition, since no explicit inverse matrix is computed, iterative methods are more suitable to on-the-fly updates from telemetry and calibration data sets. Testing and implementation will have a great impact in establishing real-time performance metrics and suitability for facility-size MOAO systems and in anchoring simulations to a real-world demonstrator as a precursor of AO for ELTs.

2. The OL control of the DMs also calls for novel temporal filtering approaches. Section 3.4 shows that the DM stability and repeatability (go-to error) is suitable for OL control, but this can be further enhanced by improving the model of the DM deformation (Guzmán et al. 2008). The application of minimum-variance techniques makes perfect applicative sense, and much insight has already been gained in using these techniques in standard closed-loop AO. They can potentially further reduce the servo-lag errors by embedding a full description of the delays (integer or fractional multiple of the sensing frame rate), the WFS and DM temporal dynamics, and the spatiotemporal properties of the disturbances (atmospheric phase, wind-shake, vibrations, and non-common-path aberrations).

The RAVEN team is currently tackling these issues, in the prospect of enhancing overall performance to increase sky coverage for full exploitation of RAVEN's capabilities.

### 4.3. WFS Simulations

The Andor iXon X3 860 camera, which uses a  $128 \times 128$  pixel EMCCD, was selected as the WFS detector for the following reasons: this camera has low read noise, can be read out at rates up to 500 frames per second, and is an affordable choice for a MOAO demonstrator. The drawback of the camera is the number of pixels. OL WFSs need a high dynamic range. Choices had to be made regarding the order of the AO system, the WFS FOV, and the pixel scale, and their impacts on performance and sky coverage had to be understood.

#### 4.3.1. System Order

A critical system design parameter is the order of the MOAO system (i.e., the number of subapertures across the WFS and

number of actuators across the diameter of the DM). Three major factors were important:

1. A commitment to producing an instrument that will deliver science data from IRCS to astronomers has been made, so sky coverage is a driving concern. Larger subapertures will allow the WFSs to work with fainter stars (as shown in § 3.5).
2. The system should not be limited in performance by fitting error if it can be avoided.
3. A tertiary concern is that if the number of subapertures were large, the number of pixels per subaperture will be limited.

A number of simulations were performed that evaluated the performance with different numbers of subapertures. The simulations were performed for the case of the NGS + LGS implementation of RAVEN (Fig. 7) and show that the RAVEN performance requirements can be met with a system order of 10. The gains achieved as a function of system order are more modest for the NGS-only implementation of RAVEN. In that

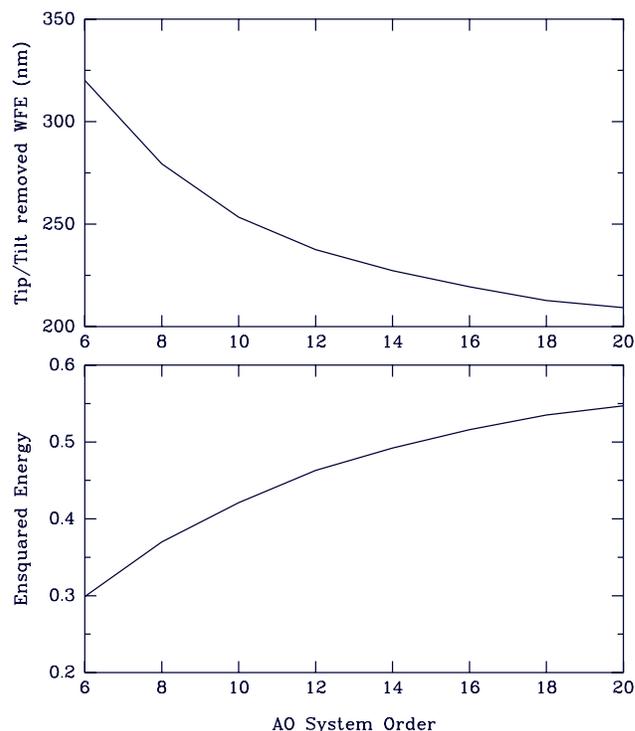


FIG. 7.—Tip/tilt-removed WFE (*top*) and ensquared energy in the  $H$  band (*bottom*) vs. the system order of RAVEN using three NGSs (on a  $45''$  diameter ring) and on-axis LGS. Median WFEs and EEs of field points within  $45''$  of the field center are used. Implementation and WFS noise are included in these simulations. We find that increasing the system order (the number of subapertures across the WFS and number of actuators across the diameter of the DM) will decrease the WFE and increase the EE, but RAVEN will more than meet the performance requirements with a system order of 10 while still allowing for a relatively large collecting area in each subaperture and a higher sky coverage.

case, tomographic error dominates the error budget, so as long as the system order remains greater than  $\sim 6$ , RAVEN performance will not significantly decrease.

4.3.2. WFS FOV

Once the system order had been chosen, the optimal WFS FOV could be determined. We simulated a single OL WFS for different atmospheres with  $r_0 = 7, 10, \text{ and } 15 \text{ cm}$  at a wavelength of  $500 \text{ nm}$  and an outer scale of  $L_0 = 100 \text{ m}$ . Using uncorrelated realizations of the atmosphere, we evaluated the cumulative distribution of source photons within each of the  $d_0 = 0.8 \text{ m}$  sub-apertures. If the IQ is very poor ( $r_0 = 7 \text{ cm}$ ), one needs a WFS with a FOV approaching  $5''$  to not miss light. Under conditions closer to median, a  $\sim 3''$  FOV could suffice.

We also simulated the effect of different outer scales on the required FOV, but found only a very weak dependence. For RAVEN, we chose a FOV of  $4.8''$ . This should give us some margin if effects other than the atmosphere shift the spots on the WFS.

4.3.3. Plate Scale

Having chosen the system order and set limits on the FOV, the effects of different WFS plate scales on RAVEN performance were examined. To preserve a  $\sim 5''$  WFS FOV, the plate scale needs to be between  $0.4 \text{ and } 0.5'' \text{ pixel}^{-1}$ . At this scale the seeing-limited WFS spots are undersampled, and the concern arises that the sensitivity of the WFS may be compromised. However, simulations indicated that WFE due to the undersampled PSF was only  $72 \text{ nm}$  for  $0.4''$  pixels. This centroiding accuracy was achieved using a MF centroiding algorithm. It is unlikely, however, that an unmodified MF algorithm will be used for centroiding in the RAVEN RTC, because the MF is linear over a limited FOV. We expect to employ either a modified matched-filter algorithm with a moving center defined by the center of gravity or a correlation centroiding algorithm in the RAVEN RTC. We are studying open-loop centroiding in more detail so that errors arising from undersampling the PSF are minimized.

TABLE 6  
H-BAND PERFORMANCE FOR DIFFERENT ATMOSPHERIC PROFILES

Subaru IQ profile	LGS + NGS		NGS only	
	EE <sub>140</sub>	Strehl	EE <sub>140</sub>	Strehl
25% .....	0.59	0.48	0.53	0.35
50% .....	0.51	0.45	0.43	0.29
75% .....	0.39	0.35	0.30	0.18

4.4. Performance Using Different Atmospheric Profiles

The performance of the RAVEN AO architecture was benchmarked using the Subaru 50% profile, but the performance of the baseline system was also checked using the Subaru 25% and 75% atmospheric profiles (Table 1). Simulated RAVEN performance for the different profiles, assuming no WFS noise, are presented in Table 6. Even when  $r_0$  is small, RAVEN will still concentrate a significant fraction of the light within the 140 mas IRCS slit.

While the 30% ensquared-energy requirement (under median conditions) will not be met with RAVEN when the IQ is poor, it appears that the performance will gracefully degrade as conditions worsen. If conditions improve, the tomographic error will still dominate the WFE budget. If the central LGS is used, high Strehl ratios ( $\sim 50\%$ ) over a  $1.5'$  diameter field can potentially be measured.

4.5. Performance as a Function of Zenith Angle

MOAO performance will depend on the zenith angle of the science target, of course. The distance from each of the atmospheric layers to the telescope is stretched by the air mass, which is equivalent to  $AM \equiv \sec(\gamma)$ , and the Fried parameter,  $r_0$ , is proportional to  $\sec(\gamma)^{-3/5}$ . For the LGS, the distance of the generated beacon and the thickness of the sodium layer are also proportional to  $\sec(\gamma)$ . As a result, the LGS dims in proportion to  $\cos(\gamma)$  if we neglect other contributions to LGS

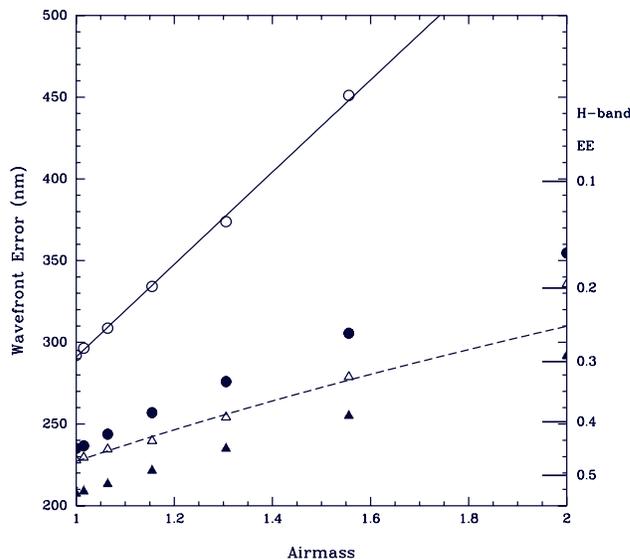


FIG. 8.—Wavefront error (open circles; scale on left) and H-band EE (filled circles; scale on right) as a function of air mass for a point near a NGS (triangles) and  $45''$  from the nearest NGS (circles). The performance drops more rapidly for points in the field far from NGSs. At the location of a NGS, the WFE increases as  $AM^{1/2}$ .

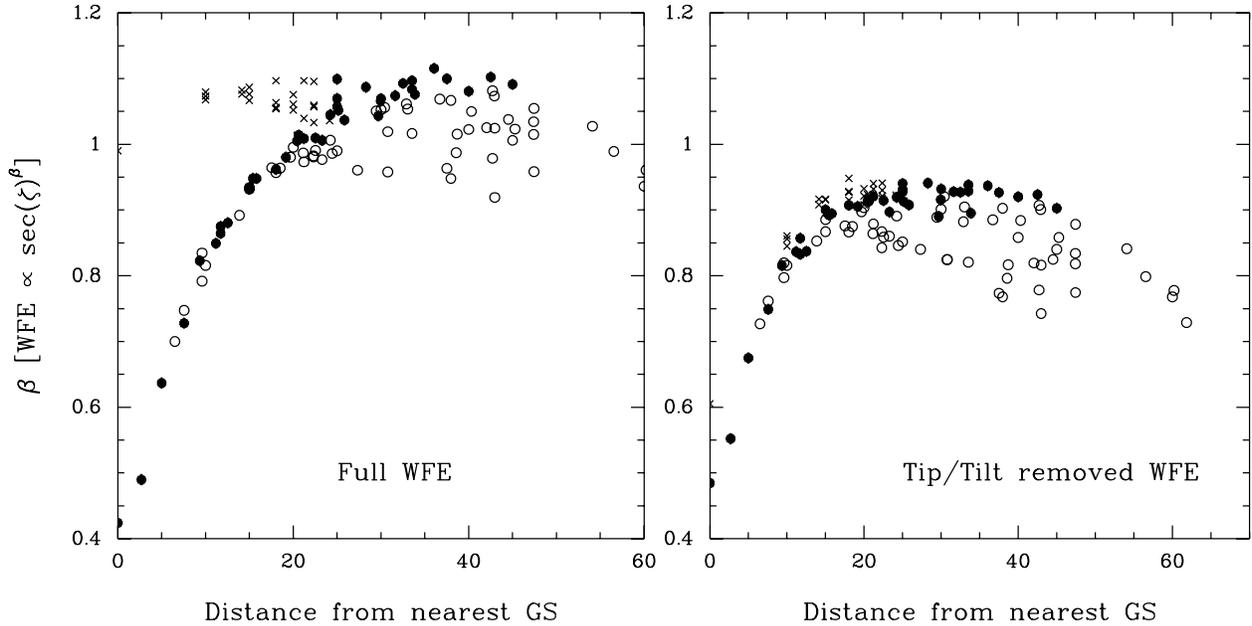


FIG. 9.—Power-law index  $\beta$ , as defined in the relation  $WFE \propto \sec(\zeta)^\beta$ , vs. distance to the nearest guide star considering the full WFE (*left*) and WFE with the tip/tilt component removed (*right*). Points are labeled by whether the closest guide star is the LGS (*x*), points are inside the 45'' diameter ring on which the NGSS sit (*filled circles*) or are outside that ring (*open circles*).

brightness, including fasor power and polarization and the Earth’s magnetic field (Holzlöhner et al. 2010).

Figure 8 shows the wavefront error and Strehl ratio at the center of the FOR for the baseline RAVEN configuration without the central LGS as a function of air mass. The derived WFEs increase rapidly and are well fit by a power law. Even though the generalized fitting error is a function of the square root of air mass (§ 3.4), we found that the best-fit exponent for a power-law function was greater. One can understand this by considering the overlapping footprints of the guide stars in the metapupils corresponding to different atmospheric layers as the zenith angle changes. As the zenith angle increases, the layers essentially get farther from the telescope, and the metapupils of the guide stars separate, leaving a larger portion of the metapupil for a given layer less well-sampled or even unsensed. If one looks at the power-law index as a function of distance from the nearest guide star, this effect becomes clear (Fig. 9).

For classical AO systems dominated by fitting error, the WFE should be proportional to  $AM^{1/2}$ . Since a large fraction of the RAVEN WFE budget is dominated by tomographic error, we expect that the power-law index  $\beta$ , defined from the relation  $WFE \propto AM^\beta$ , will vary with distance from the guide stars (Fig. 9). We see that near the NGSS,  $\beta \sim 0.5$ , but that  $\beta$  rises to  $\sim 1$  away from the NGSS. The power-law index  $\beta$  is large even very close to the LGS, because atmospheric turbulence is dominated by tip/tilt, which is not sensed by the LGS WFS. If tip/tilt

is excluded, the relation between distance to the guide star and  $\beta$  is virtually indistinguishable between LGSs and NGSS.

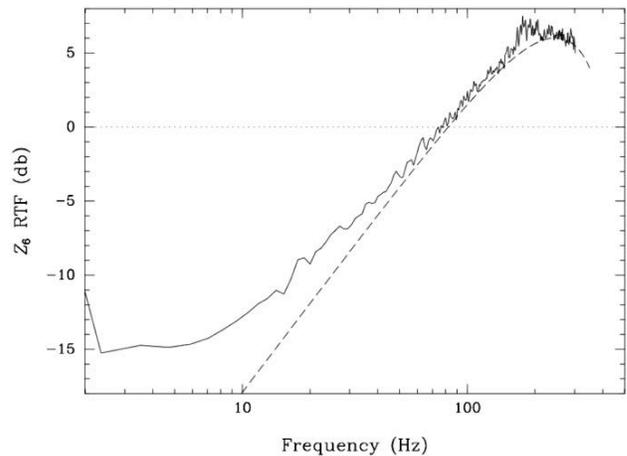


FIG. 10.—VOLT rejection transfer function. The behavior of the RTF at frequencies greater than 40 Hz matches the theoretical OL RTF (*dashed line*), but at low frequencies, the open-loop VOLT system was not making the full rejection. This same behavior is observed with both ViLLaGEs and CANARY.

#### 4.6. Sampling Frequency

ViLLaGEs, VOLT, and CANARY have all demonstrated that open-loop correction is possible. The gain of an open-loop system can be set to unity, because there is no feedback; every wavefront measurement made is applied to the DM. The overshoot is limited to a factor of 2 (because there is no feedback) at one-fourth of the sampling frequency,  $f_s$  (assuming a one-frame delay). The correction bandwidth is high; for VOLT running with  $f_s = 750$  Hz, the correction bandwidth was approximately 70 Hz (Fig. 10). However, it was also noticed that the RTF flattened out at low temporal frequencies (seemingly independently of spatial frequency). Our best rejection was approximately  $-15$  dB with VOLT (CANARY was able to achieve  $\sim -25$  dB of rejection; E. Gendron 2011, private communication). This is most likely due to some misalignment in the system, which we hope to simulate, but it seems unlikely that this type of error can ever be completely eliminated in RAVEN. Much was learned about alignment and calibration from VOLT, and significant improvement on the OL RTF is anticipated.

To determine the effect that this flattened RTF would have on RAVEN performance, an OL RTF with and without this plateau was simulated. Assumptions made include median image quality of  $r_0 = 15.6$  cm at  $\lambda = 500$  nm, an infinite outer scale, and a wind speed of  $8 \text{ m s}^{-1}$ . The residual wavefront error in Zernike modes 7–45 (excluding tip/tilt and second-order radial modes, because errors on these modes will not lead to significant losses in ensquared energy) was measured for different sampling frequencies between  $50 \text{ Hz} < f_s < 500 \text{ Hz}$ , different RTF saturation levels of  $-15$  dB and  $-25$  dB, and no saturation (Table 7).

The results are quite independent of sampling frequency, at least above 100 Hz. If RAVEN can be run with  $f_s = 100$  Hz, approximately a half-magnitude more can be gained in the limiting magnitude (down to  $m_R = 15$  for the NGS-only case), which will in turn have a big impact on the RAVEN sky coverage. It is also noted that saturation levels must be kept at or below  $-15$  dB; at this level, it is already becoming a significant source of WFE. If RAVEN can be kept aligned and calibrated so that this saturation level is approximately  $-25$  dB, the open-loop temporal error will be only a minor contribution to the total WFE budget.

#### 5. SUMMARY AND FUTURE DIRECTIONS

A baseline MOAO system architecture for RAVEN has been established and the expected performance of such a system has been simulated using two independent modeling tools, MAOS and OMAO. These two independently developed AO simulation tools give excellent agreement for the expected performance of RAVEN. Based on these results, it has been established that RAVEN should be able to meet the design requirement that 30% ensquared energy be delivered within a 140 mas wide IRCS slit if three NGSs are used, and perfor-

TABLE 7  
RAVEN OPEN-LOOP TEMPORAL WFE

Saturation level	Sampling frequency				
	50	100	180	250	500
No saturation	91(103)	49(56)	31(36)	25(28)	17(19)
$-25$ dB	91(105)	52(62)	36(48)	32(45)	28(43)
$-15$ dB	108(147)	90(134)	88(133)	88(133)	88(133)

NOTES.—Measured in nm rms, excluding first- and second-order Zernike modes. Results in parentheses are residual wavefront errors, excluding only tip and tilt.

mance will improve dramatically if the single Subaru facility on-axis LGS WFS is also included (up to  $\sim 40\%$  ensquared energy in median conditions). Employing one LGS beacon also greatly improves the sky coverage, because RAVEN can operate with the LGS in conjunction with just two NGSs that are  $m_R < 14.5$  and perhaps even fainter.

A broad spectrum of the system parameter space has been explored. We have looked at RAVEN performance as a function of guide star magnitude. We expect RAVEN to be able to provide the required ensquared energy if the NGSs have magnitude  $R < 14.5$  ( $R < 15$  if the LGS is also used). Trade studies on the DMs and WFSs have allowed us to settle on an AO system architecture that includes order  $10 \times 10$  WFSs with a  $4.8''$  FOV and a pixel scale of  $0.4'' \text{ pixel}^{-1}$  and an order  $11 \times 11$  DM. We explored how the performance changes as the asterism diameter decreases, zenith angle increases, and different atmospheric profiles were used. We found that the degradation of performance with zenith angle is dependent on the distance to the nearest guide star; the farther away the science object sits from a guide star, the more rapidly performance degrades. A study of the performance with sampling frequency using a realistic open-loop rejection transfer function that incorporates the effects of optical misalignments and imperfect calibration concluded that little performance loss is expected if the sampling frequency is greater than or equal to 180 Hz (if the open-loop rejection transfer function plateaus at  $-15$  dB at low temporal frequencies, one has larger wavefront errors, of course, but one can use  $f_s = 100$  Hz without additional WFEs).

As the RAVEN project progresses, we are placing a high priority on developing tools to simulate RAVEN alignment and calibration. Exploration of these simulations will improve the understanding of the spatial and temporal frequencies affected by open-loop misalignments. Existing open-loop on-sky experiments reported RTFs that flattened-off at low frequencies. Another issue that needs further study is open-loop centroiding. The MF seems to work quite well under the conditions studied, but concern about a steep dropoff in performance must be addressed in the event that the spots move off of the MF. Thresholded center of gravity should work well in the high-S/N regime, but the sensitivity of the thresholded center of

gravity may be limited, due to the undersampled WFS spots. In addition to these algorithms, we will also explore how well suited correlation centroiding and a modified MF (with a floating center set by the center of gravity) are to open-loop centroiding and RAVEN. Finally, we will continue to explore the feasibility

of using advanced control and reconstruction algorithms in the context of RAVEN. We hope that RAVEN will help demonstrate that MOAO projects are indeed feasible and that MOAO instruments are capable of delivering significant multiplex advantages over single-target IFSs.

## REFERENCES

- Ammons, S. M., Gavel, D. T., Dillon, D. R., Reinig, M., Grigsby, B., & Morzinski, K. M. 2008, *Proc. SPIE*, 7015, 701546
- Ammons, S. M., et al. 2010, *PASP*, 122, 573
- Andersen, D. R., Fischer, M., Conan, R., Fletcher, M., & Véran, J.-P. 2008, *Proc. SPIE*, 7015, 70150H
- Andersen, D. R., Fischer, M., & Véran, J.-P. 2009, *OSA Technical Digest, AOT4A4*
- Andersen, D., et al. 2006, *Proc. SPIE*, 6269, 62694K
- Costille, A., et al. 2010, *J. Opt. Soc. Am. A*, 27, 469
- Cuby, J.-G., et al. 2010, *Proc. SPIE*, 7735, 77352D
- Ellerbroek, B. L., Gilles, L., & Vogel, C. R. 2003, *Proc. SPIE*, 4839, 989
- Els, S., et al. 2009, *PASP*, 121, 527
- Flicker, R., Rigaut, F. J., & Ellerbroek, B. L. 2000, *Proc. SPIE*, 4007, 1032
- Gavel, D., Bauman, B., Dekany, R., Britton, M., & Andersen, D. 2006, *Proc. SPIE*, 6272, 62720R
- Gavel, D., et al. 2008, *Proc. SPIE*, 7015, 70150G
- Gendron, E., et al. 2010, *Proc. SPIE*, 7736, 77360P
- Gendron, E., et al. 2011, *A&A*, 529, L2
- Gilles, L., & Ellerbroek, B. 2006, *Appl. Opt.*, 45, 6568
- Gilles, L., & Ellerbroek, B. 2008, *J. Opt. Soc. Am. A*, 25, 2427
- Guzmán, D., et al. 2008, *Proc. SPIE*, 7015, 70153X
- Hammer, F., et al. 2002, in *Scientific Drivers for ESO Future VLT/VLTI Instrumentation*, ed. J. Bergeron, & G. Monnet (Berlin: Springer), 139
- Holzlohner, R., et al. 2010, *Proc. SPIE*, 7736, 77360V
- Jackson, K., Conan, R., & Véran, J.-P. 2010, *Proc. SPIE*, 7736, 77364K
- Johnston, D. C., & Welsh, B. M. 1991, *Proc. SPIE*, 1542, 76
- Kissler-Patig, M. 2010, in *1st Adaptive Optics for Extremely Large Telescopes Conf. (Les Ulis: EDP Sciences)*, 01001
- Martinez, P., et al. 2010, *Messenger*, 141, 5
- Minowa, Y., et al. 2010, *Proc. SPIE*, 7736, 77363N
- Miyashita, A., et al. 2004, *Proc. SPIE*, 5489, 207
- Morris, T., et al. 2010 in *1st Adaptive Optics for Extremely Large Telescopes Conf. (Les Ulis: EDP Sciences)*, 08003
- Morzinski, K., et al. 2010, *Proc. SPIE*, 7736, 77361O
- Nelson, J. 2008, *Proc. SPIE*, 6986, 698602
- Primmerman, C. A., Murphy, D. V., Page, D. A., Zollars, B. G., & Barclay, H. T. 1991, *Nature*, 353, 141
- Puech, M., et al. 2006, *NewA Rev.*, 50, 382
- Ragazzoni, R. 1999, in *ESO Conf. Proc. 56, Astronomy with Adaptive Optics*, ed. D. Bonaccini (Garching: ESO), 651
- Ragazzoni, R., Marchetti, E., & Rigaut, F. 1999, *A&A*, 342, L53
- Rigaut, F. J., Ellerbroek, B. L., & Flicker, R. 2000, *Proc. SPIE*, 4007, 1022
- Robin, A. C., Reylé, Derrière, S., & Picaud, S. 2003, *A&A*, 409, 523 (erratum 416, 157 [2004])
- Shectman, S., & Johns, M. 2010, *Proc. SPIE*, 7733, 77331Y
- Tokovinin, A., Le Louarn, M., Viard, E., Hubin, N., & Conan, R. 2001, *A&A*, 378, 710
- Tokunaga, A., et al. 1998, *Proc. SPIE*, 3354, 512
- Vidal, F., et al. 2010, in *1st Adaptive Optics for Extremely Large Telescopes Conf. (Les Ulis: EDP Sciences)*, E7001V
- Wizinowich, P., et al. 2010, *Proc. SPIE*, 7736, 77360K

## Appendix B

Static and predictive tomographic reconstruction for wide-field multi-object adaptive optics systems

# Static and predictive tomographic reconstruction for wide-field multi-object adaptive optics systems

C. Correia,<sup>1,2,\*</sup> K. Jackson,<sup>2</sup> J.-P. Véran,<sup>3</sup> D. Andersen,<sup>2</sup> O. Lardière,<sup>2</sup> and C. Bradley<sup>2</sup>

<sup>1</sup>Centre for Astrophysics, University of Porto, Rua das Estrelas, 4150-762 Porto, Portugal

<sup>2</sup>Adaptive Optics Laboratory, University of Victoria, Victoria, British Columbia V8P 5C2, Canada

<sup>3</sup>National Research Council, Herzberg Institute of Astrophysics, 5071 West Saanich Road, Victoria, British Columbia V9E 2E7, Canada

\*Corresponding author: ccorreia@astro.up.pt

Received July 31, 2013; revised October 28, 2013; accepted November 9, 2013;  
posted November 19, 2013 (Doc. ID 194920); published December 11, 2013

Multi-object adaptive optics (MOAO) systems are still in their infancy: their complex optical designs for tomographic, wide-field wavefront sensing, coupled with open-loop (OL) correction, make their calibration a challenge. The correction of a discrete number of specific directions in the field allows for streamlined application of a general class of spatio-angular algorithms, initially proposed in Whiteley *et al.* [J. Opt. Soc. Am. A **15**, 2097 (1998)], which is compatible with partial on-line calibration. The recent Learn & Apply algorithm from Vidal *et al.* [J. Opt. Soc. Am. A **27**, A253 (2010)] can then be reinterpreted in a broader framework of tomographic algorithms and is shown to be a special case that exploits the particulars of OL and aperture-plane phase conjugation. An extension to embed a temporal prediction step to tackle sky-coverage limitations is discussed. The trade-off between lengthening the camera integration period, therefore increasing system lag error, and the resulting improvement in SNR can be shifted to higher guide-star magnitudes by introducing temporal prediction. The derivation of the optimal predictor and a comparison to suboptimal autoregressive models is provided using temporal structure functions. It is shown using end-to-end simulations of Raven, the MOAO science, and technology demonstrator for the 8 m Subaru telescope that prediction allows by itself the use of 1-magnitude-fainter guide stars. © 2013 Optical Society of America

OCIS codes: (010.1080) Active or adaptive optics; (010.1330) Atmospheric turbulence.

<http://dx.doi.org/10.1364/JOSAA.31.000101>

## 1. MULTI-OBJECT ADAPTIVE OPTICS

Single-object adaptive optics (AO) systems are now routinely used in most major observatories to attenuate the blurring effect of the Earth's atmosphere when imaging through turbulence. This blurring results in loss of angular resolution, set by the ratio of the imaging wavelength  $\lambda$  and the telescope diameter. Classical AO systems use a single wavefront sensor (WFS) and deformable mirror (DM) driven in real-time to measure and correct the wavefront phase aberrations in a single direction.

To achieve correction over a large field, angular anisoplanatism must be overcome. This phenomenon is linked to the variation of the optical disturbances across the field and is characterized by the anisoplanatic angle,  $\theta_0 \propto \lambda^{6/5}$ , typically of the order of tens of arcseconds in the H band [1].

Two approaches are envisioned for enlarging the isoplanatic angle: multi-object AO (MOAO) and multi-conjugate AO (MCAO). Both involve tomographic estimation of the 3D atmospheric wavefront disturbance using information from multiple WFSs locked on multiple guide stars, which probe different lines of sight through the atmosphere. The two approaches differ in how the correction is applied:

- In MCAO, one places multiple DMs in a series, each optically conjugated to a different atmospheric altitude. Due to the finite number of DMs, such systems will still suffer from generalized anisoplanatism [2].

- In MOAO, after the information from multiple WFSs is combined into a tomographic estimate of the turbulence [3], multiple science pick-off arms are placed on the scientifically interesting targets in the field. Each science channel contains a DM, which makes the optimal turbulence correction in its science direction. This parallel approach promises to increase the field over which AO corrections can be applied to 5 arcmin or even 10 arcmin.

However, unlike classical and MCAO closed loop AO systems, in which the WFSs sit after the DMs, MOAO systems require open-loop (OL) estimation of the atmosphere over a large field but only a few discrete number of correction directions. As such, the WFSs do not capture the DM shape, being bound to measure the total uncorrected disturbance (as opposed to measuring only the residual). The DM shape is not fed back to the system, being blindly applied to the device.

Because of its OL aspect, such systems stress the need for accurate calibration and call for (quasi) model-independent reconstructors. One solution for calibrating MOAO systems is the so-called Learn & Apply algorithm [4] in which the tomographic reconstructor is formulated as a minimum mean square error (MMSE) optimization problem, whose solution involves covariance matrices that can be directly measured in the MOAO system. These matrices encode a certain number of instrumental systematics and are, therefore, preferable to synthetic versions that may undermodel the actual phenomena in the real system.

A related MMSE solution had been previously proposed in Whiteley *et al.* [5] extending results of Valley [6] when dealing with optimal compensation of angular anisoplanatism in classical AO. As it turns out, MOAO can be seen as a generalization to multiple directions and any number of ground-conjugated correction directions. Therefore, a streamlined application of Whiteley's algorithm is formulated. Furthermore, the static Strehl-optimal reconstructors under consideration for 30–40 m class telescopes' closed-loop MCAO systems [7] can be worked out as variations (albeit mathematically equivalent) of the MMSE reconstructor.

As it stands, only partial computation of the tomographic reconstructor can be done from the data. The remainder, performing the angular extrapolation in the field from the guide stars to the science object direction(s), relies still on fitting an *a priori* atmospheric model to the data and cannot be circumvented.

Furthermore, to overcome temporal loop delays, or at least to mitigate them, the static reconstructors are then extended with an extra temporal prediction step. Potential predictive autoregressive (AR) models are compared to the one-step spatio-angular (SA) temporal solution under a common temporal structure-function analysis framework.

Temporal prediction has the potential to relax the temporal constraints, thus allowing for larger integration times and/or real-time processing. This may in turn increase the limiting magnitude of stars that can be used for guiding.

The performance analysis is done for Raven, the MOAO science and technology demonstrator for the 8 m Subaru telescope, supported by Monte Carlo simulations and experimental results from the Raven optical bench.

This paper has three main sections: in Section 2, two equivalent formulations are recalled and their features outlined. The explicit layered tomographic reconstructor allows more easily the formulation of an intermediate predictive step, whose options are discussed in Section 3. Finally, in Section 4, the gain in terms of limiting magnitude for the Raven demonstrator are presented and compared to numerical simulations.

## 2. SPATIAL, STREHL-OPTIMAL, STATIC RECONSTRUCTORS

Under the hypothesis that the turbulent atmosphere is a sum of  $N_l$  thin layers located in a discrete number of different altitudes  $h_j$ , the aperture-plane phase  $\omega(\boldsymbol{\rho}, \boldsymbol{\theta}, t)$  indexed by the bi-dimensional spatial coordinate vector  $\boldsymbol{\rho} = (\rho_x, \rho_y)$  in direction  $\boldsymbol{\theta} = (\theta_x, \theta_y)$  at time  $t$  is defined as

$$\omega(\boldsymbol{\rho}, \boldsymbol{\theta}, t) = \sum_{j=1}^{N_l} \mathcal{W}_j(\boldsymbol{\rho} + h_j \boldsymbol{\theta}, t), \quad (1)$$

where  $\mathcal{W}_j(\boldsymbol{\rho}, t)$  is the  $j$ th-layer wave-front.

In the following, assume the phase is conveniently expanded onto a piston-removed Zernike orthonormal modal basis [8] with a finite number of  $N_z$  polynomials,

$$\omega(\boldsymbol{\rho}, \boldsymbol{\theta}, t) = \sum_{i=2}^{N_z} \phi_i(\boldsymbol{\theta}, t) \mathbf{Z}_i(\boldsymbol{\rho}/R), \quad (2)$$

with modal coefficients

$$\phi_i(\boldsymbol{\theta}, t) = \frac{1}{R^2} \iint \mathbf{Z}_i(\boldsymbol{\rho}/R) \omega(\boldsymbol{\rho}, \boldsymbol{\theta}, t) \Omega(\boldsymbol{\rho}) d^2 \boldsymbol{\rho}, \quad (3)$$

where  $\Omega(\boldsymbol{\rho})$  is the aperture function,  $R$  the telescope primary mirror radius, and  $\mathbf{Z}_i(\boldsymbol{\rho}/R)$  the  $i$ th polynomial function.

Using matrix formulation, the resulting aperture-plane wavefront coefficients  $\boldsymbol{\phi}(\boldsymbol{\theta}, t)$  in the near-field approximation relate to the coefficients  $\boldsymbol{\varphi}$  of the wavefront phase defined over a discrete number of layers in the volume by a simple matrix multiplication [9]:

$$\boldsymbol{\phi}(\boldsymbol{\theta}, t) = \mathbf{P}_\theta \boldsymbol{\varphi}(t), \quad (4)$$

where

$$\boldsymbol{\varphi} \triangleq \left[ \boldsymbol{\varphi}_0^T \quad \cdots \quad \boldsymbol{\varphi}_{N_l}^T \right]^T \quad (5)$$

is a concatenation of phase coefficients of the decomposition of  $\mathcal{W}_j(\boldsymbol{\rho}, t)$ ,  $1 \leq j \leq N_l$  onto the basis function set, and  $\mathbf{P}_\theta$  is a cookie-cutter matrix that remaps and sums the Zernike coefficients from the wavefront modal expansion in the meta-pupils to the pupil-plane. The superscript  $T$  represents vector transpose.

As is common in AO, the aperture-plane phase is not measured directly. Instead, the WF is estimated (reconstructed) from some measured (and noisy) data that is statistically related to it. Consider only the widely used Hartmann–Shack wavefront sensor (HS-WFS), which provides the phase gradient  $\boldsymbol{s} = (s_x, s_y)$  with noise.

In MOAO, the objective cost function is the minimization of the aperture-plane residual phase variance (MV) for individual science directions  $\boldsymbol{\beta}_i \in \mathfrak{R}^{1 \times 2}$ :

$$\mathbf{E} = \arg \min_{\mathbf{E}} (\|\hat{\boldsymbol{\phi}}(\boldsymbol{\beta}_i) - \hat{\boldsymbol{\phi}}(\boldsymbol{\beta}_i)\|_{L_2(\Omega)}^2), \quad (6)$$

where  $\hat{\boldsymbol{\phi}}$  is the actual phase (or its coefficients as is the case here),  $\hat{\boldsymbol{\phi}}(\boldsymbol{\beta}_i) \triangleq \mathbf{E} \boldsymbol{s}$  is the estimated phase (estimated quantities are indicated by the hat symbol),  $\boldsymbol{s}$  are noisy measurements,  $L_2$  is the Euclidean norm over the aperture  $\Omega$ , and  $\langle \cdot \cdot \cdot \rangle$  is the ensemble average over time for an individual optimization direction. As is shown in [10], the minimization is independent of the direction  $\boldsymbol{\beta}_i$ , i.e., it is equivalent to minimize the residual phase variance over a given field-of-view (FoV) of interest. Due to the particular optimization for individual directions in MOAO, the reconstructors can be simplified, and a more compact solution is found as is described next.

### A. Equivalent Tomographic Reconstructors

Let the following forward-measurement model, which makes explicit use of the layered phase vector [9,10],

$$\boldsymbol{s}_\alpha(t) = \boldsymbol{\Gamma} \mathbf{P}_\alpha \boldsymbol{\varphi}(t) + \boldsymbol{\eta}(t) \quad (7)$$

with  $\boldsymbol{s}_\alpha(t) \in \mathfrak{R}^{(N_\alpha \times N_s) \times 1}$  a column vector of  $N_\alpha \times N_s$  measurements for all the  $N_\alpha$  GS directions, obtained using the pupil-plane gradient operator  $\boldsymbol{\Gamma} \triangleq \text{diag}\{\boldsymbol{\Gamma}_1, \dots, \boldsymbol{\Gamma}_{N_\alpha}\} \in \mathfrak{R}^{N_\alpha N_s \times N_\alpha N_z}$  that concatenates  $N_\alpha$  gradient operators for each individual WFS.  $\boldsymbol{\Gamma}_i$  relates the  $N_z$  phase dimensions in the aperture in the directions  $\alpha \in \mathfrak{R}^{N_\alpha \times 2}$  to the  $N_s$  measurements. Noise is represented by  $\boldsymbol{\eta}(t) \triangleq \in \mathfrak{R}^{N_\alpha N_s \times 1}$ .

The MV reconstructor providing the aperture-plane phase estimate in the  $\beta$  science directions is obtained by minimizing the partial derivatives of Eq. (6) with respect to  $E'$  [7,10,11]:

$$\hat{\phi}_\beta = \mathbf{P}_\beta \mathbf{E} s_\alpha, \quad (8a)$$

$$= \mathbf{P}_\beta \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^T \rangle \mathbf{P}_\alpha^T \Gamma^T (\mathbf{G} \mathbf{P}_\alpha \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^T \rangle \mathbf{P}_\alpha^T \Gamma^T + \langle \boldsymbol{\eta} \boldsymbol{\eta}^T \rangle)^{-1} s_\alpha, \quad (8b)$$

where  $\boldsymbol{\phi}_\beta \in \mathfrak{R}^{N_\beta N_z \times 1}$  is a vector of aperture-plane phase coefficients from the decomposition of the WF in all the  $N_\beta$  science directions  $\beta \in \mathfrak{R}^{N_\beta \times 2}$ . This representation is preferred in closed loop wide-field AO systems and has been extensively used previously coupled to pseudo-OL control [12]. It is particularly suited to multi-conjugate AO, where the turbulence profile is needed prior to fitting to a DM's influence functions for correction. The deterministic fitting step is a least-squares fit to the DM influence functions that does not depend upon the measurement noise nor turbulence statistics.

With this formulation, the turbulence profile is explicitly estimated in the layers before being collapsed to the pupil plane through  $\mathbf{P}_\beta$ ; it is therefore referred to as “explicit tomography reconstructor” in the remainder of this paper.

Before continuing, note that Eq. (8) is equivalent to the MMSE solution with a simplified measurement model involving the pupil-plane turbulence only:

$$s_\alpha(t) = \mathbf{G} \boldsymbol{\phi}_\alpha(t) + \boldsymbol{\eta}(t). \quad (9)$$

Assuming  $s$  and  $\boldsymbol{\phi}$  are zero-mean and jointly Gaussian, direct application of the MMSE solution to estimate the aperture-plane phase in the  $N_\beta$ -science directions of interest yields [13]

$$\mathcal{E}\{\boldsymbol{\phi}_\beta | s_\alpha\} \triangleq \Sigma_{(\boldsymbol{\phi}_\beta, s_\alpha)} \Sigma_{s_\alpha}^{-1} s_\alpha = \hat{\boldsymbol{\phi}}_\beta, \quad (10)$$

where  $\mathcal{E}\{X|Y\}$  stands for the mathematical expectation of  $X$  conditioned to  $Y$ . Since in general  $\beta \neq \alpha$ , Eq. (10) follows from  $\mathcal{E}\{\boldsymbol{\phi}_\beta | s_\alpha\} = \mathcal{E}\{\boldsymbol{\phi}_\beta | \mathcal{E}\{\boldsymbol{\phi}_\alpha | s_\alpha\}\}$ . Given that the conditioning relates only to the last available measurement (as opposed to present and previous measurements), these reconstructors are labeled as *static*. Developing terms in Eq. (10) using Eq. (9), the reconstructor becomes

$$\hat{\boldsymbol{\phi}}_\beta \triangleq \langle \boldsymbol{\phi}_\beta \boldsymbol{\phi}_\alpha^T \rangle \Gamma^T (\mathbf{G} \langle \boldsymbol{\phi}_\alpha \boldsymbol{\phi}_\alpha^T \rangle \mathbf{G}^T + \langle \boldsymbol{\eta} \boldsymbol{\eta}^T \rangle)^{-1} s_\alpha. \quad (11)$$

Equation (11) converts to Eq. (8) by setting

$$\langle \boldsymbol{\phi}_\beta \boldsymbol{\phi}_\alpha^T \rangle = \mathbf{P}_\beta \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^T \rangle \mathbf{P}_\alpha^T, \quad (12a)$$

$$\langle \boldsymbol{\phi}_\alpha \boldsymbol{\phi}_\alpha^T \rangle = \mathbf{P}_\alpha \langle \boldsymbol{\varphi} \boldsymbol{\varphi}^T \rangle \mathbf{P}_\alpha^T. \quad (12b)$$

This MMSE reconstructor is dubbed SA reconstructor, a term coined by Rodolphe Conan, to the best of the authors' knowledge, on account of the nature of the covariance matrices involved in its definition. It can be seen as a generalization of the work of Whiteley *et al.* [5] in seeking the optimal anisoplanatic reconstructor in classical AO to the tomographic, multiple sensor case. It has several convenient features: it is much faster to compute off-line than the explicit tomography reconstructor and circumvents the truncated expansion on a modal basis.

The numerical equivalence of Eq. (12) is further explored in Section 4.A for the Raven demonstrator [1].

When phase is expanded onto the Zernike polynomials, the SA cross-covariance functions can be analytically computed for the infinite outer-scale case of turbulence [14,15]:

$$\langle \phi_i(0) \phi_j(\xi) \rangle = 3.895 \left( \frac{D}{r_0} \right)^{\frac{5}{3}} \frac{\int_0^{h_{\max}} C_n^2(h) I_{ij} \left( \frac{\xi h}{R} \right) dh}{\int_0^\infty C_n^2(h) dh} \quad (13)$$

with  $D = 2R$ , the telescope diameter;  $r_0$ , the Fried parameter;  $h$ , the altitude;  $\xi$ , the angle between the pupils over which the Zernike polynomials are defined;  $C_n^2$ , the atmospheric vertical profile; and  $I_{ij}(x)$ , a term involving 1D numerical integration. Equation (13) has been extended for the finite outer-scale case in [16] and later extensively used and generalized in [5,17]. The layered spatial covariance matrix  $\langle \boldsymbol{\varphi} \boldsymbol{\varphi}^T \rangle$  from Eq. (8b) is a block-diagonal matrix (layers are independent) and can be found in [8] for the infinite outer-scale case and in [18] for the finite case.

## B. Spatio Angular versus Explicit Tomography

OL operation of MOAO systems poses a complex problem for system calibration. MOAO is exposed to potential issues, such as misregistration, field-dependent distortions, and irregular sensitivity (to cite a few) since the WFSs do not “see” the actual DM figures. In closed-loop systems, the recursive nature of the loop using feedback allows for partial compensation of miscalibrations.

The SA formulation is thus particularly amenable to MOAO systems since, at least in principle, the covariance matrices composing the reconstructor can be directly computed from acquired data. This could not be achieved in a closed-loop system. In regular on-sky operation,  $\Sigma_{s_\alpha}$  from Eq. (10) can be recorded on-line (with use of some caution to ensure statistical convergence and proper acquisition). This covariance matrix provides a wealth of valuable information about the system itself with the strong potential to carry the signature of misalignments and other spatially variant discrepancies not taken into account in the models. Also, and equally important, it carries an imprint of the vertical turbulence profile, from which the  $C_n^2$  profile and integrated atmospheric parameters estimation, such as  $r_0$ ,  $L_0$ , and eventually the wind velocity vectors, can be estimated through data postprocessing using the built-in SLODAR method [19]. These parameters are then used to constrain a model for the second covariance matrix.

The Learn & Apply algorithm from [4] can now be presented as the SA reconstructor with covariances defined in slope space instead of phase space. The choice for remaining in slope space is due to system calibration; the algorithm directly collects measurements' covariances and makes the angular extrapolation in measurement space, bypassing an explicit slopes-to-phase reconstruction. Thus the projection onto and from phase space is circumvented at the expense of having larger covariance matrices (there are roughly twice as many slopes as estimated phase vectors).

Provided access to the layered turbulence is granted, linear prediction could be plugged into the explicit tomography reconstructor as an intermediary step to counter intrinsic temporal delays in the AO system, from integration and processing. The OL working environment excludes dynamical stability issues associated to feedback systems. Using Taylor's

frozen-flow approximation, the wavefront on each layer can be predicted before collapsing it onto the pupil-plane in the directions of interest. Options are more broadly investigated in Section 3.

### 3. PREDICTIVE MODELS FOR ATMOSPHERIC RECONSTRUCTION

As is well known for any AO system, the availability of sufficiently bright guide-stars overshadows the utility of AO. Therefore the sky coverage, i.e., the percentage of available sky for observing, imposes a strong constraint on the number of observable targets. Sky coverage can be improved by either making better use of every single photon (better wavefront sensing) or by relaxing the temporal lag-error constraint, which allows for longer integration times if the error can be partly overcome by temporal prediction. Lag error is intrinsic to any AO system due to the discrete nature of the measurements, their processing and correction. All combined requires several milliseconds to read-out and compute which stresses the importance of a predictive capability. The lag error definition adopted here is the sum of the integration time  $T_s$  plus the pure-delay  $\Delta$  error encompassing the real-time processing of measurements and computation of DM commands. See the temporal diagram in Fig. 1.

When  $\Delta$  is not an integer multiple of the frame-rate measurements, commands become asynchronous. Working in OL greatly simplifies the implementation when compared to closed-loop systems since in the measurement model the asynchronous DM commands do not intervene. The latter are computed as the weighted average over two consecutive phase instances spanned by the DM commands, using the fractional delays as weights. Such procedure would considerably increase the computational complexity, in particular when prediction is on. For this reason, in the remainder the case,  $\Delta \neq 0$  is considered in the simulations but is not taken into account on the reconstructors.

#### A. Beating Down Temporal Lag Errors

As previously stated, the explicit-tomography formulation is well suited to temporal prediction. Note, though, that with some extra complexity, the same could be built into the SA formulation [5], albeit with no flexibility as to the choice of the predictive model.

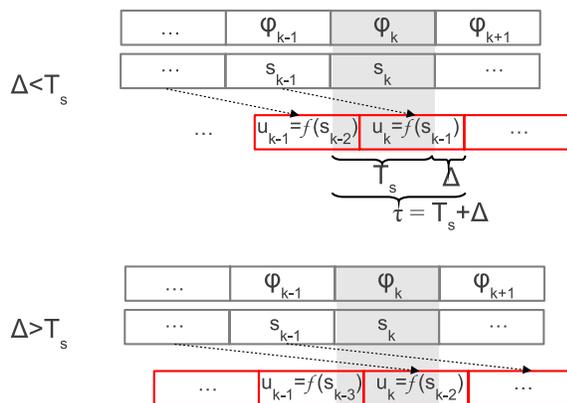


Fig. 1. Temporal diagrams. Top:  $\text{Int}(\Delta/T_s) = 0$  commands  $u_k$  are conditioned to measurement  $s_{k-1}$ . Bottom:  $T_s < \Delta < 2T_s$ ;  $u_k$  is conditioned to  $s_{k-2}$ .

#### Proposition

Temporal prediction can be achieved by means of a linear operation  $\mathcal{P}$  on current and past samples of the phase vectors, such that

$$\hat{\phi}_{\beta_j, k+1} = \mathbf{P}_{\beta_j} \mathcal{P}[\hat{\phi}_k, \hat{\phi}_{k-1}, \dots, \hat{\phi}_{k-n}], \quad (14)$$

where  $\mathcal{P}$  is a linear prediction operator. For the case of no prediction,  $\mathcal{P} = \mathbf{I}$ .

In the following, Taylor's frozen-flow hypothesis is assumed:

$$\omega(\rho, t + \tau) = \omega(\rho + v \cdot \tau, t). \quad (15)$$

A first approach uses the fact that the spatial gradient of phase is measured by the HS-WFS. As a consequence, with a truncated polynomial expansion of Eq. (15) to the first-order [20], turbulence could be predicted. However, this model is restricted to single-conjugated AO systems and lacks generality for wide-field tomographic phase estimation—unless layer-oriented tomography is used [21], a framework not adopted here.

#### B. One-Step SA Predictor

A more general method that complies with tomographic phase estimation is sought. Time evolution of expansion coefficients leads to non-null cross correlations [17], which should be fully grasped to obtain an optimal prediction. Figure 2 depicts the modal cross-correlation functions for tip and focus (polynomials 2 and 4, respectively, following [8]) up to the 9th mode recurring to Eq. (13).

Recasting the problem as a proper criterion minimization, the best linear predictor is the solution to the following criterion:

$$\mathcal{A}_\tau = \arg \min_{\mathcal{A}_\tau} \langle \|\varphi(t + \tau) - \mathcal{A}_\tau \varphi(t)\|_{L_2(\Omega)}^2 \rangle, \quad (16)$$

yielding

$$\mathcal{A}_\tau \triangleq \langle \varphi_{k+1} \varphi_k^T \rangle \langle \varphi_k \varphi_k^T \rangle^{-1}. \quad (17)$$

This operation in what follows is called a “one-step SA predictor” (the face-on pattern of  $\mathcal{A}_\tau$  is depicted in Fig. 3). It is a nondiagonal, densely populated matrix; a sign of the temporal cross-correlations between modes coming into play. The same predictor is outlined by [11] for the case of phase represented by its samples on a regular grid of points (also called the zonal representation).

The one-step cross-covariance matrix  $\langle \varphi_{k+1} \varphi_k^T \rangle$  can be easily estimated from Eq. (13), under the assumption of isotropic and stationary turbulence. However, it has weak dependence on the wind direction (see, for instance, the short-term auto-correlation for modes tip and tilt in Fig. 4). Therefore, for pure frozen flow, a modal approach can be short of a high-fidelity phase-shifting approach, although this could be done at the expense of extra computation.

Furthermore, Eq. (17) is a general method to generate and predict phase in a 2D plane (any wind velocity can be used), according to the Markovian model [13]:

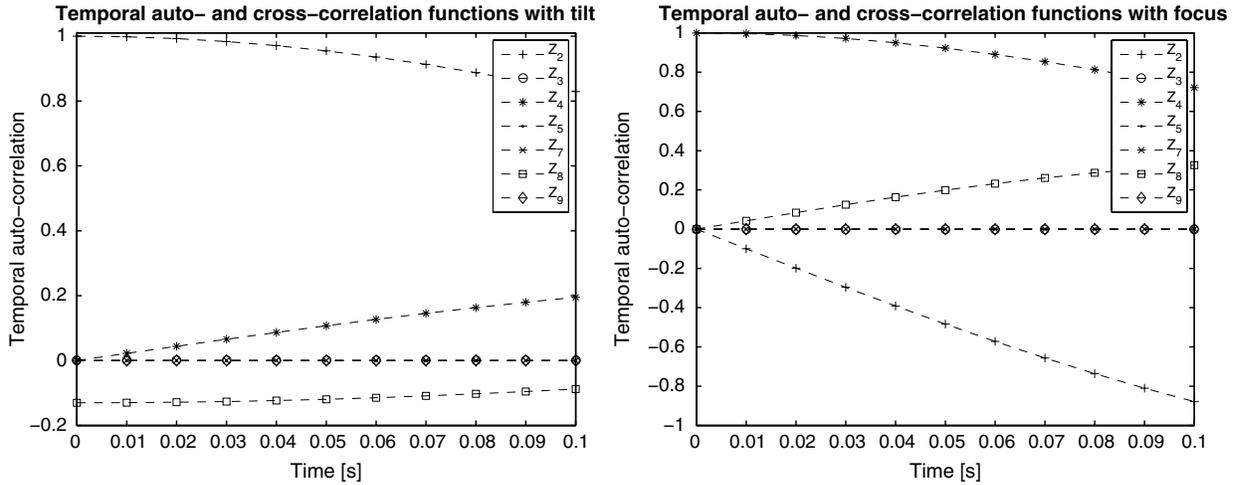


Fig. 2. Temporal auto- and cross-correlation functions for the tip and focus modes. Although at  $t = 0$ , these plots show the spatial values in [8], for  $t > 0$  correlations appear and vanish as shown. One can see, for the case of focus, that although it isn't correlated to any other mode plotted for  $t = 0$ , a strong (anti-) correlation appears with tilt as time elapses.

$$\boldsymbol{\varphi}_{k+1} = \mathcal{A}_\tau \boldsymbol{\varphi}_k + \boldsymbol{\varepsilon}_k^\delta, \quad (18)$$

where  $\boldsymbol{\varepsilon}_k^\delta$  is an excitation noise whose properties are fixed to guarantee proper turbulence statistics.

This model can either be used for generating a fully developed turbulence or for control-oriented purposes such as a Kalman filter gain computation [13]. In either case, the pair  $(\mathcal{A}_\tau, \Sigma_\varepsilon^\delta)$  is required. The excitation noise covariance matrix  $\Sigma_\varepsilon^\delta$  is found from imposing the output statistics to be those of a Kolmogorov or von-Kármán model. Hence  $\Sigma_\varepsilon^\delta = \Sigma_\varphi - \mathcal{A}_\tau \Sigma_\varphi \mathcal{A}_\tau^T$  since  $\langle \boldsymbol{\varphi}_{k+1} \boldsymbol{\varphi}_{k+1}^T \rangle = \mathcal{A}_\tau \Sigma_\varphi \mathcal{A}_\tau^T + \Sigma_\varepsilon^\delta$ . Note  $\Sigma_\varphi \triangleq \langle \boldsymbol{\varphi}_{k+1} \boldsymbol{\varphi}_{k+1}^T \rangle \triangleq \langle \boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^T \rangle$ , which, due to stationarity, loses its temporal dependence.

A further remark is that with such Markovian processes (future values are only conditioned to present ones), a simplification can be done. It consists of considering the total delay  $\tau = T_s + \Delta m$  simplicity as the sampling step of the predictive model. The asynchronous case is thus dealt with efficiently with no recourse to weighted averaging.

Assémat's method for simulating infinitely long, nonstationary phase screens [22] turns out to be a truncated version of

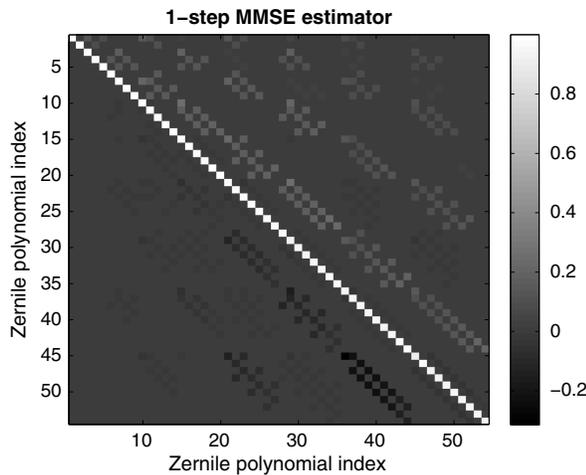


Fig. 3. Face-on pattern of  $\mathcal{A}_\tau$ , single-layer, computed for nine radial orders ( $N_z = 55$ ) with  $|v| = 15 \text{ ms}^{-1}$  and  $\tau = 10 \text{ ms}$ , from which results a displacement of  $0.15 \text{ m}$ .

Eq. (17) when a pointwise (zonal) representation of the phase is used. However, only a subset of the columns of  $\mathcal{A}_\tau$  is used to include bounded region correlations. Extensions can be found in [23,24].

Despite the temporal cross-correlation being taken into account in this model,  $\mathcal{A}_\tau$  is strongly diagonally dominated, suggesting that simpler diagonal models, i.e., mode-by-mode, could be potentially applied, and indeed they have been extensively used in AO simulations in the form of AR models [25]. Their features and relation to the SA predictor are explored next.

### C. Autoregressive Models

For several AO applications constrained by real-time computational burden, using simpler diagonal AR models can be quite appealing as they circumvent  $\mathcal{A}_\tau$ , being a dense matrix. Although these relatively coarse models are not adapted to simulating atmospheric turbulence, they are used instead for prediction when embedded in the reconstructor and plugged into the controller—as is done with Kalman filtering [26,27]—for off-line computation of optimal gains.

An AR model of order  $n$  is defined by the recursion

$$\boldsymbol{\varphi}_{k+1} = f(\boldsymbol{\varphi}_k, \dots, \boldsymbol{\varphi}_{k-n+1}) + \boldsymbol{\varepsilon}_k, \quad (19)$$

where  $f(\dots)$  is a linear function yet to be defined, and  $\boldsymbol{\varepsilon}_k$  is a Gaussian-distributed spectrally white zero-mean random sequence with variance such that the output variance is conserved as in the previous section.

An AR model of the first order (AR1) is simply

$$\boldsymbol{\varphi}_{k+1}^{\text{AR1}} = \mathbf{A}_{\text{AR1}} \boldsymbol{\varphi}_k^{\text{AR1}} + \boldsymbol{\varepsilon}_k^{\text{AR1}}, \quad (20)$$

where a diagonal  $\mathbf{A}_{\text{AR1}}$  replaces  $\mathcal{A}_\tau$  in Eq. (18). The excitation noise covariance matrix for the model in Eq. (20) is likewise computed as in Section 3.B.

For a second-order model (AR2):

$$\boldsymbol{\varphi}_{k+1}^{\text{AR2}} = \mathbf{A}_{\text{AR2}} \boldsymbol{\varphi}_k^{\text{AR2}} + \mathbf{B}_{\text{AR2}} \boldsymbol{\varphi}_{k-1}^{\text{AR2}} + \boldsymbol{\varepsilon}_k^{\text{AR2}}. \quad (21)$$

Imposing  $\Sigma_\varphi^{\text{AR2}} = \Sigma_\varphi$  as before,  $\Sigma_\varepsilon^{\text{AR2}}$  is found from

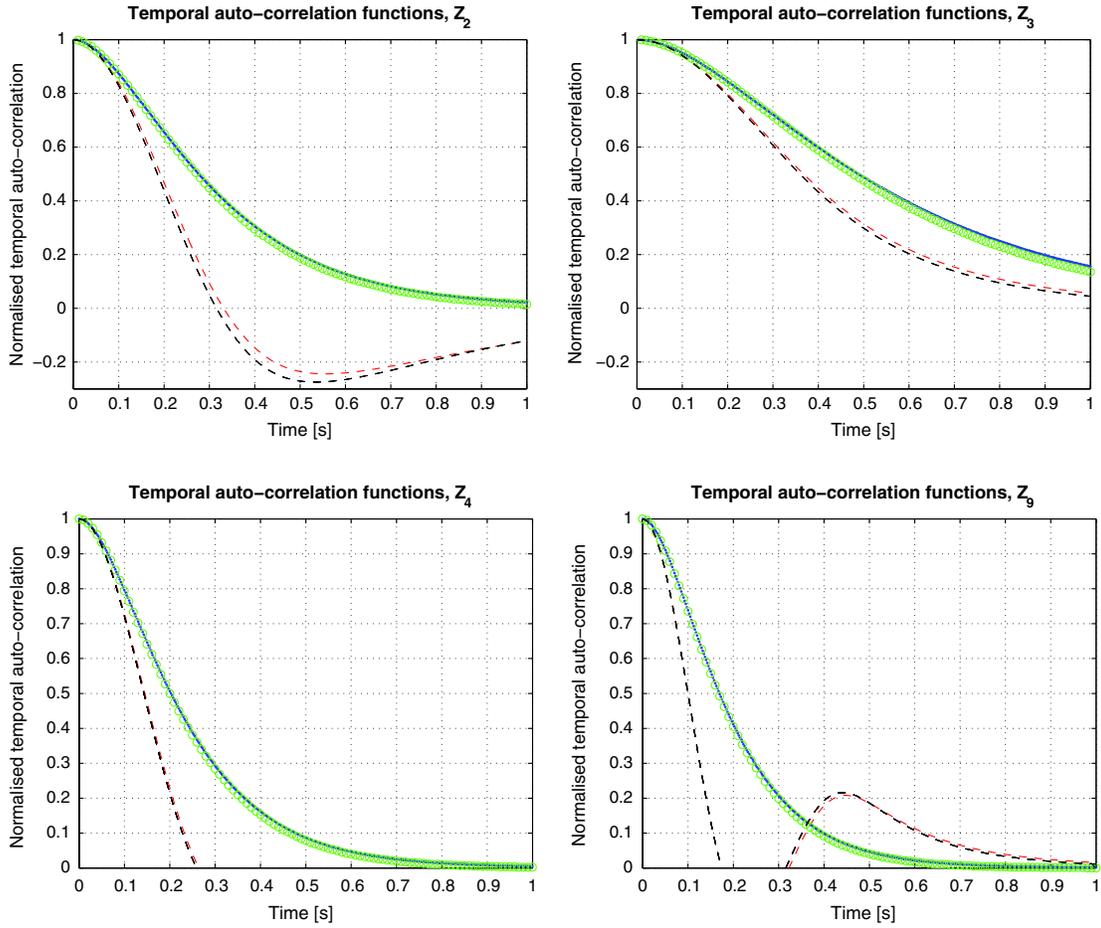


Fig. 4. Theoretical temporal auto-correlation functions assuming frozen flow computed directly (black dash) or as the Fourier-transformed temporal PSDs (red dash). Comparison against the second-order continuous (blue dots) and discrete (green circles) predictive models fitting the initial 50 ms of the theoretical curves. Model fitting uses the Broyden–Fletcher–Goldfarb–Shanno method.

$$\begin{aligned} \Sigma_e^{\text{AR}2} &= \Sigma_\varphi - \mathbf{A}_{\text{AR}2} \Sigma_\varphi \mathbf{A}_{\text{AR}2}^T - \mathbf{B}_{\text{AR}2} \Sigma_\varphi \mathbf{B}_{\text{AR}2}^T - \mathbf{A}_{\text{AR}2} \Sigma_{1\tau}^{\text{AR}2} \mathbf{B}_{\text{AR}2}^T \\ &\quad - \mathbf{B}_{\text{AR}2} \Sigma_{1\tau}^{\text{AR}2} \mathbf{A}_{\text{AR}2}^T, \end{aligned} \quad (22a)$$

$$\Sigma_{1\tau}^{\text{AR}2} = \langle \varphi_{k+1}^{\text{AR}2} (\varphi_k^{\text{AR}2})^T \rangle \triangleq \langle \varphi_k^{\text{AR}2} (\varphi_{k-1}^{\text{AR}2})^T \rangle = (\mathbf{I} - \mathbf{B}_{\text{AR}2})^{-1} \mathbf{A}_{\text{AR}2} \Sigma_\varphi, \quad (22b)$$

where  $\Sigma_{1\tau}^{\text{AR}2}$  is a one-step modal spatiotemporal cross-covariance matrix.

Similarly, for a third-order model (AR3):

$$\varphi_{k+1}^{\text{AR}3} = \mathbf{A}_{\text{AR}3} \varphi_k^{\text{AR}3} + \mathbf{B}_{\text{AR}3} \varphi_{k-1}^{\text{AR}3} + \mathbf{C}_{\text{AR}3} \varphi_{k-2}^{\text{AR}3} + \varepsilon_k^{\text{AR}3}, \quad (23)$$

the noise covariance matrix is found from

$$\begin{aligned} \Sigma_e^{\text{AR}3} &= \Sigma_\varphi - \mathbf{A}_{\text{AR}3} \Sigma_\varphi \mathbf{A}_{\text{AR}3}^T - \mathbf{B}_{\text{AR}3} \Sigma_\varphi \mathbf{B}_{\text{AR}3}^T - \mathbf{C}_{\text{AR}3} \Sigma_\varphi \mathbf{C}_{\text{AR}3}^T \\ &\quad - \mathbf{A}_{\text{AR}3} \Sigma_{1\tau}^{\text{AR}3} \mathbf{B}_{\text{AR}3}^T - \mathbf{B}_{\text{AR}3} \Sigma_{1\tau}^{\text{AR}3} \mathbf{A}_{\text{AR}3}^T - \mathbf{B}_{\text{AR}3} \Sigma_{1\tau}^{\text{AR}3} \mathbf{C}_{\text{AR}3}^T \\ &\quad - \mathbf{C}_{\text{AR}3} \Sigma_{1\tau}^{\text{AR}3} \mathbf{B}_{\text{AR}3}^T - \mathbf{A}_{\text{AR}3} \Sigma_{2\tau}^{\text{AR}3} \mathbf{C}_{\text{AR}3}^T - \mathbf{C}_{\text{AR}3} \Sigma_{2\tau}^{\text{AR}3} \mathbf{A}_{\text{AR}3}^T, \end{aligned} \quad (24a)$$

where  $\Sigma_{1\tau}^{\text{AR}3}$  is again the one-step modal cross-covariance matrix; this time for the AR3 model, and  $\Sigma_{2\tau}^{\text{AR}3} \triangleq \langle \varphi_{k+1}^{\text{AR}3}$

$(\varphi_{k-1}^{\text{AR}3})^T \rangle \triangleq \langle \varphi_k^{\text{AR}3} (\varphi_{k-2}^{\text{AR}3})^T \rangle$  is a two-step modal spatiotemporal cross-covariance matrix, with

$$\Sigma_{1\tau}^{\text{AR}3} = (\mathbf{B}_{\text{AR}3} + \mathbf{C}_{\text{AR}3} \mathbf{A}_{\text{AR}3} + \mathbf{C}_{\text{AR}3}^2)^{-1} (\mathbf{A}_{\text{AR}3} + \mathbf{C}_{\text{AR}3} \mathbf{B}_{\text{AR}3}) \Sigma_\varphi, \quad (24b)$$

$$\Sigma_{2\tau}^{\text{AR}3} = \mathbf{B}_{\text{AR}3} \Sigma_\varphi + (\mathbf{A}_{\text{AR}3} + \mathbf{C}_{\text{AR}3} \mathbf{B}_{\text{AR}3}) \Sigma_{1\tau}^{\text{AR}3}. \quad (24c)$$

### 1. Model Identification Using Zernike Polynomials

A practical method to identify matrices  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  is to fit the initial  $T_{\text{fit}}$  seconds of the temporal auto-correlation function of each and every mode as previously suggested in [28]. An alternative is to match the decorrelation at the coherence time of each mode [29]. For AR1 models, both strategies lead to roughly the same modal decorrelation functions and can therefore be used interchangeably. For higher-order models (AR2 and AR3), fitting the initial  $T_f$  seconds leads to the overall best prediction performances, although the existence of a general  $T_{\text{fit}}$  is debatable and subject to optimization.

The temporal auto-correlation function is the Fourier-transformed temporal power spectral density (PSD), a direct application of the Wiener–Khinchine theorem (on a mode-by-mode basis). The PSDs are computed from [30].

A much faster and effective method builds on the SA covariance matrices of Eq. (13), with an angle given by  $\xi = \tau \times |v|/h$  (under Taylor's frozen-flow hypothesis). This method is much more cost-effective since the correlation functions are computed directly and only over the  $T_{\text{fit}}$  horizon and not beyond, leading to considerable computational gains.

Figure 4 plots the temporal auto-correlation functions from Zernike polynomials 1 to 9 using the Wiener–Kinchine theorem against the SA formulation. The numerical agreement is remarkable, with slight discrepancies due to the limits of integration in either case. Also over-plotted are the second-order continuous model and its AR discretization for  $T_s = 10$  ms, when fitting the initial  $T_{\text{fit}} = 50$  ms. The AR (discrete time) model is obtained using standard zero-order-hold discretization procedures.

#### D. Prediction Error Structure Functions

The predictive capabilities of different models can be assessed using the predicted phase temporal structure function. Define the temporal lag error (no prediction) as the average phase variance taken two instants apart, i.e., the temporal structure function of the phase,

$$D_t(\tau) = \langle |\omega(\boldsymbol{\rho}, t) - \omega(\boldsymbol{\rho}, t + \tau)|^2 \rangle, \quad (25)$$

which relates to the spatial structure function of phase  $D_\rho$  through Taylor's frozen-flow hypothesis by

$$D_t(\tau) = \langle |\omega(\boldsymbol{\rho}, t) - \omega(\boldsymbol{\rho} - \mathbf{v} \cdot \boldsymbol{\tau}, t)|^2 \rangle_\rho = D_\rho(v \cdot \boldsymbol{\tau}), \quad (26)$$

with  $\langle \dots \rangle_\rho$  the ensemble average over the spatial distances  $\rho = |\boldsymbol{\rho}|$ ,  $v = |\mathbf{v}|$ . Expanding the squared term and assuming stationarity gives  $2(C(0) - C_t(\tau)) = 2(C(0) - C_\rho(v\tau))$ , where  $C_t(\cdot)$  and  $C_\rho(\cdot)$  are the phase temporal and spatial covariance functions.  $C_t(\tau) = C_\rho(v\tau)$ , i.e., the temporal covariance function is deduced from the phase spatial covariance function evaluated at  $\rho = v\tau$ .

The phase spatial structure function is given by [6]

$$D_\rho(\rho) = \left( \frac{L_0}{r_0} \right)^{5/3} \times \frac{2^{1/6} \Gamma(11/6)}{\pi^{8/3}} \left[ \frac{24}{5} \Gamma\left(\frac{6}{5}\right) \right]^{5/6} \times \left[ \frac{\Gamma(5/6)}{2^{1/6}} - \left( \frac{2\pi\rho}{L_0} \right)^{5/6} K_{5/6} \left( \frac{2\pi\rho}{L_0} \right) \right], \quad (27)$$

with  $L_0$  the outer scale of turbulence,  $r_0$  Fried's coherence length,  $\Gamma$  the "gamma" function, and  $K_{5/6}$  a modified Bessel function of the third kind.

It is useful to consider the general formulation for the temporal lag error that is compatible with the case of atmospheric prediction:

$$\sigma_{\text{lag}}^2(\tau, p) = \langle \|\mathbf{P}_\theta(\boldsymbol{\varphi}_k - \hat{\boldsymbol{\varphi}}_k)\|_{L_2(\Omega)}^2 \rangle, \quad (28)$$

where  $\hat{\boldsymbol{\varphi}}_k$  is the phase estimate using any of the predictive models of order  $p = \{0, 1, 2, 3\}$ .

In the **no** prediction case, the estimated phase is simply a replication of the phase at the previous time step,  $\hat{\boldsymbol{\varphi}}_k = \boldsymbol{\varphi}_{k-1}$ . The temporal lag error from Eq. (28) becomes

$$\sigma_{\text{lag}}^2(\tau, p = 0) = 2 \text{trace}\{\mathbf{P}_\theta(\Sigma_\varphi - \Sigma_{1\tau})\mathbf{P}_\theta^T\} \triangleq D_t(\tau), \quad (29)$$

which is the temporal structure function of phase from Eq. (26), with the one-step covariance matrix  $\Sigma_{1\tau} = \Sigma(\rho = v\tau)$  computed from Eq. (13) with a proper angle.

These temporal structure functions can now be expanded for the case of predicted phase. For first-order models (the AR1 or the one-step SA predictor) one has  $\hat{\boldsymbol{\varphi}}_k = \mathbf{A}\boldsymbol{\varphi}_{k-1}$  (note the excitation noise is not included). Developing Eq. (28) yields

$$\sigma_{\text{lag}}^2(\tau, p = 1) = \text{trace}\{\mathbf{P}_\theta(\Sigma_\varphi + \mathbf{A}\Sigma_\varphi\mathbf{A}^T - 2\Sigma_{1\tau}\mathbf{A}^T)\mathbf{P}_\theta^T\}, \quad (30)$$

with  $\mathbf{A} = \mathbf{A}_{\text{AR1}}$  for the AR1 model and  $\mathbf{A} = \mathbf{A}_\tau$  for the SA predictor model. For the AR2:

$$\sigma_{\text{lag}}^2(\tau, p = 2) = \text{trace}\{\mathbf{P}_\theta(\Sigma_\varphi + \mathbf{A}_{\text{AR2}}\Sigma_\varphi\mathbf{A}_{\text{AR2}} + \mathbf{B}_{\text{AR2}}\Sigma_\varphi\mathbf{B}_{\text{AR2}} - 2\mathbf{A}_{\text{AR2}}\Sigma_{1\tau} + \mathbf{A}_{\text{AR2}}\Sigma_{1\tau}\mathbf{B}_{\text{AR2}}^T - 2\mathbf{B}_{\text{AR2}}\Sigma_{2\tau})\mathbf{P}_\theta^T\}, \quad (31)$$

with  $\Sigma_{1\tau} = \Sigma(\rho = v\tau)$  and  $\Sigma_{2\tau} = \Sigma(\rho = 2v\tau)$ .

Likewise for the AR3:

$$\sigma_{\text{lag}}^2(\tau, p = 3) = \text{trace}\{\mathbf{P}_\theta(\Sigma_\varphi + \mathbf{A}_{\text{AR3}}\Sigma_\varphi\mathbf{A}_{\text{AR3}} + \mathbf{B}_{\text{AR3}}\Sigma_\varphi\mathbf{B}_{\text{AR3}} + \mathbf{C}\Sigma_\varphi\mathbf{C}_{\text{AR3}} - 2\mathbf{A}_{\text{AR3}}\Sigma_{1\tau} - 2\mathbf{B}_{\text{AR3}}\Sigma_{2\tau} - 2\mathbf{C}_{\text{AR3}}\Sigma_{3\tau} + \mathbf{A}_{\text{AR3}}\Sigma_{1\tau}\mathbf{B}_{\text{AR3}}^T + \mathbf{B}_{\text{AR3}}\Sigma_{1\tau}\mathbf{C}_{\text{AR3}}^T + \mathbf{A}_{\text{AR3}}\Sigma_{2\tau}\mathbf{C}_{\text{AR3}}^T)\mathbf{P}_\theta^T\}, \quad (32)$$

with  $\Sigma_{3\tau} = \Sigma(\rho = 3v\tau)$ . The matrices  $\Sigma_{p\tau}$  are computed for a fully developed turbulence following Eq. (13) and thus differ from those associated with any specific model.

In Fig. 5, the theoretical lag errors (for the cases **no** prediction and **with** prediction) are plotted [31]. Several comments follow:

- The one-step SA predictor provides the best performance (as expected) for large lags above  $\sim 5$  ms, which stems from its optimality [it minimizes the prediction error variance in Eq. (16)] and cross-mode prediction. However, the predictor is only a truncated version; in practice, a finite number of modes is to be used. When more modes are added in, the one-step SA predictor beats the AR2-3 models for lower lags.

- For lags below  $\sim 5$  ms AR models of orders 2 and 3 slightly outperform the one-step SA predictor since the short term decorrelation of the AR models is quite similar to that of the turbulence.

- The AR2 is a sufficient model-order above which the performance gains are little and not worth the increased complexity for both parameter identification or real-time processing; an AR2 model presents a temporal PSD  $\propto \nu^{-4}$ , which is a good approximation to the actual phase whose spectrum is  $\propto \nu^{-11/3}$ , with  $\nu = |\nu|$  the temporal frequency vector modulus.

- Although the AR1 has been successfully used in control-oriented models for Kalman filtering [25,29,32], it performs quite poorly to predict phase and is patently incapable to provide any improvement over the **no** prediction case. Simulations show that when the off-diagonal values of the one-step SA predictor are nulled out, the predictive capability degrades to that of a diagonal AR1, suggesting that it is this feature that plays the most important role in the predictive process.

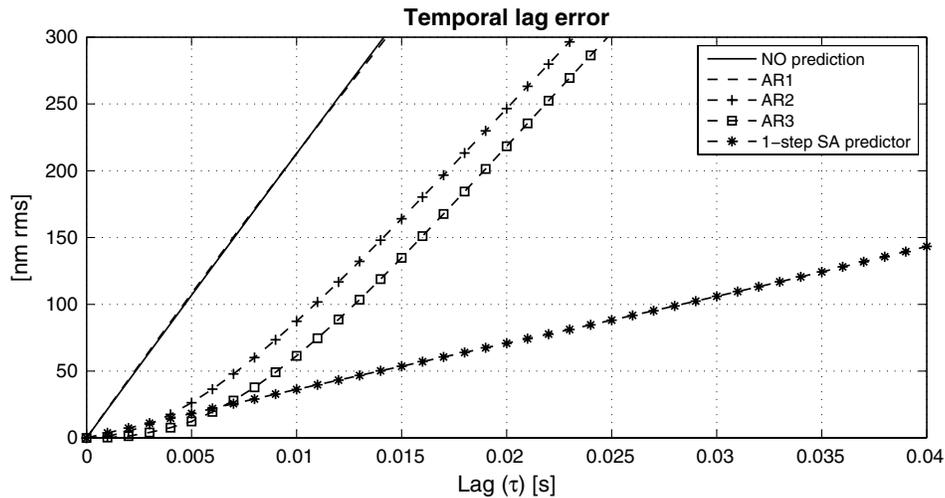


Fig. 5. Comparison of temporal lag errors on an equivalent single layer atmosphere (see Table 1 for further parameters).

Despite the results above, the offline computations of AR models in general are quite complex in fitting the auto-correlation functions (Section 3.C.1) and are much higher than the SA offline load, which does not require a fitting algorithm. Only addressing the synchronous case, the reconstructor dimensions are given by  $r = N_\beta N_a N_\alpha N_s$  for the cases of no prediction, AR1 prediction, and SA prediction. With  $p \geq 2$  one has  $r = N_\beta N_z (N_a + p N_\alpha N_s)$ . Thus the real-time computational load is not increased over a static reconstructor for AR1 and one-step SA prediction models. For the Raven case, whose parameters are given in Table 1, this corresponds to a 30% increase for the AR2 model and an 85% increase for the AR3 model.

#### 4. INCREASING THE LIMITING MAGNITUDE FOR THE RAVEN PROJECT

Raven is a MOAO science and technology demonstrator, which will be the first MOAO instrument on an 8 m class telescope feeding an AO-optimized science instrument, the Subaru InfraRed Camera and Spectrograph (IRCS). The instrument will be equipped with three natural guide star (NGS), wavefront sensors (WFS), and one laser guide star (LGS) WFS to generate a tomographic reconstruction of the atmosphere in a 3.5 arcmin field of regard (FoR) for up to two science-object directions (see Fig. 6).

The specifics of the system are given below:

- **OL-WFSs:** Three deployable pick-off mirrors can patrol the telescope focal plane and pick off three NGS within a 3.5 ft FoR. Each pick-off mirror feeds one OL WFS consisting of a  $10 \times 10$  SH-WFS ( $d = 0.8$  m) with  $12 \times 12$  pixels per subaperture and a 4.8 in. FoV per subaperture. The system's baseline detector is the E2V EMCCD camera. These devices can be set to run at very high gain. In this mode, the read noise can be made almost arbitrarily small; however, the background source plus background photon noise is increased.

- An additional OL-WFS with identical specifications is fixed on-axis in order to utilize the Subaru on-axis LGS, which will improve sky coverage and/or AO correction.

- **DMs:** There are two ALPAO DMs, one in each science channel, with  $11 \times 11$  actuators and a 25 mm clear aperture.

These will contribute a constant fitting error term of 145 nm rms at  $r_0 = 15.6$  cm.

Table 1 provides further simulation and system parameters.

#### A. Equivalence of Static Reconstructors

The equivalence of the two formulations for the static reconstructors presented in Section 2 is now established numerically. It boils down to checking the similarity between the

Table 1. Raven Baseline Configuration Parameters

<b>Telescope</b>	
D	8 m
<b>Atmosphere</b>	
$r_0$	15.6 cm
$L_0$	30 m
Zenith angle	0 deg
Fractional $r_0$	[0.596; 0.224; 0.180]
Altitudes	[0; 5; 10] km
Wind speeds	[7.5; 12.5; 15] m/s
Wind direction	[0; 0; 0] deg
Sampling	8/150 m
<b>Wavefront Sensor</b>	
RON	$0.2 e^-$
$N_{\text{NGS}}$	3
NGS radii	30 arcsec
Order	$10 \times 10$
$\theta_{\text{pix}}$	0.4 arcsec
$N_{\text{pix}}$	15
$f_{\text{sample}}$	30–200 Hz
$\lambda_{\text{WFS}}$	0.7 $\mu\text{m}$
Centroiding Algorithm Thresholded Center of Gravity	
<b>DM</b>	
Order	$11 \times 11$
Stroke	Infinite
Influence	Cubic
<b>AO loop</b>	
Pure delay	$\Delta = 3$ ms
Controlled modes	$N_z = 55$
Reconstructor	$N_p = 406$ modes
Predictor	90 modes
<b>Evaluation</b>	
$\lambda_{\text{evl}}$	1.65 $\mu\text{m}$

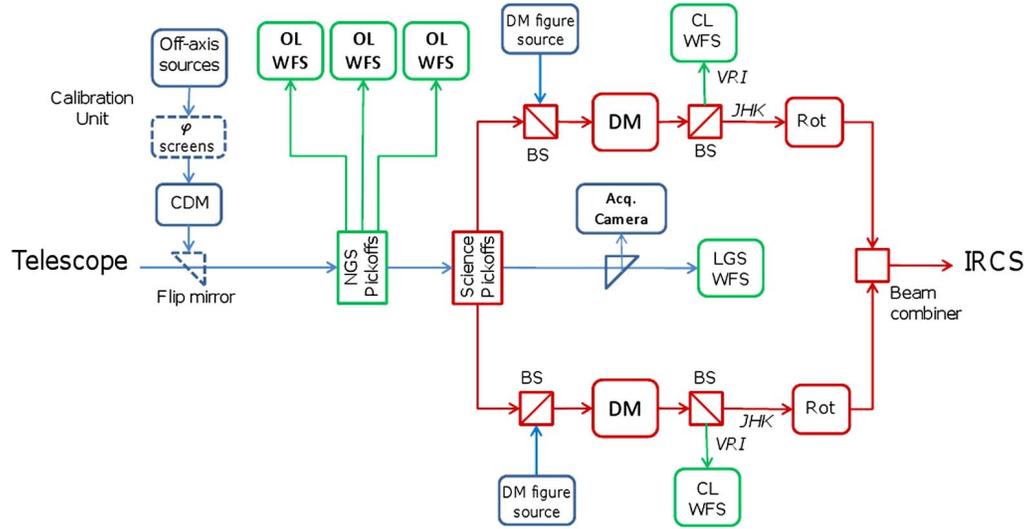


Fig. 6. Functional optical block diagram of RAVEN. Dashed blocks are deployable. Raven consists of eight main subsystems: the deployable calibration unit, the OL NGS WFSs, the science pick-offs, the science relays, the closed-loop NGS truth/figure WFSs, the beam combiner, the LGS WFS, and the acquisition camera.

angular covariance matrix between Zernike polynomials appearing in Eqs. (8b) and (11).

Their similarity increases with the increasing number of modes considered in each layer. Let  $\varphi_\beta$  be expanded onto a Zernike set of with a fixed number  $N_z$  of modes. The layered turbulence is expanded on a greater number of  $N_p$  modes.

It is instructive to assess the relative error

$$\varepsilon_i(N_p) = \frac{\text{trace}((\langle \phi_\alpha \phi_\alpha^T \rangle - \mathbf{P}_\alpha(N_p) \langle \varphi \varphi^T \rangle \mathbf{P}_\alpha(N_p)^T)}{\text{trace}(\langle \phi_\alpha \phi_\alpha^T \rangle)} \quad (33)$$

as a function of  $N_p$  to obtain a reasonable approximation to the SA covariance matrices computed using an analytical formulation. The relative error in Eq. (33) falls to the few-percent level when  $N_p$  is the result of using three times more radial orders than in  $N_z$  and below the 1% level for a factor of for, when  $N_z = 55$  (nine radial orders [8]).

With this in mind, static reconstruction of a single-shot WF from slopes has been conducted on the Raven optical bench (Fig. 6, Table 1) [1] covering the full 2 arcmin FoV. The contour plots obtained are depicted in Fig. 7, demonstrating the full equivalence of performance obtained with the SA and the explicit reconstructors using Raven bench data. Differences are in the subpercent level.

## B. Limiting Magnitude

The science gain achievable by Raven, in comparison to classical AO systems such as Subaru's AO188, will be modest because Raven will only have two science channels and provide a relatively low-order correction. Nevertheless, the 8 m aperture of the Subaru telescope enables science that is not achievable on smaller telescopes, and Raven will be capable of delivering high ensquared energy (EE) into the IRCS slit. It has been projected that, with a static tomographic reconstructor, Raven's limiting magnitude will approach 14.5 (for 30% EE) using a reduced frame rate of 180 Hz. As stated in [1], sky coverage for Raven will be low. For example, consider a point with galactic coordinates  $(b; l) = (30; 0)$ . Using the Besançon model of the galaxy [33],

one finds that there are 750 stars per square degree with  $R < 14.5$  (1040 stars per square degree with  $R < 15$ ). The probability that there are three stars with  $R < 14.5$  in a 2 arcmin diameter FOR is just 3%. This does not even account for asterisms that are unsuitable for Raven; in some cases, the science targets will not be inscribed within the potential NGS asterism, and therefore the tomographic error will be too great. Star densities are increasing as a power law at these magnitudes, so going 1 magnitude deeper can increase the density of available stars by a factor of 1.8.

Results in the previous sections indicate that the potential gains in performance can be traded for an increase in limiting magnitude. In the presence of temporal prediction, the lag error  $\sigma_{\text{lag}}^2$  decreases, and one can thus tolerate more measurement noise from fainter sources, resulting in increased noise propagation  $\sigma_{\text{np}}^2$  (averaged over the field) for the same aggregate wavefront error, defined as

$$\sigma_{\text{total}}^2 = \underbrace{\sigma_{\text{lag}}^2 + \sigma_{\text{np}}^2}_{\sigma^2} + \text{other terms}, \quad (34)$$

where other terms relates to errors that are independent from the system lag (to a first degree approximation). The latter is computed from Eq. (28) whereas the noise propagation error is

$$\sigma_{\text{np}}^2 = \text{trace}\{\mathbf{E}\Sigma_\eta\mathbf{E}^T\}, \quad (35)$$

where the noise covariance matrix  $\Sigma_\eta = \sigma_\eta^2 \mathbf{I}$  is assumed diagonal with  $\sigma_\eta^2$  the measurement noise variance on each WFS sub-aperture. The latter is a decreasing function of the star brightness, i.e., photon-noise increases for dimmer guide stars. The propagated noise  $\sigma_{\text{np}}^2$  relates to star magnitude by standard centroiding error functions [34].

The AO system bandwidth is chosen to minimize the error  $\sigma^2 = \sigma_{\text{lag}}^2 + \sigma_{\text{np}}^2$ . Thus one has

$$\sigma^2 = \sigma_{\text{np}}^2(m_v, \tau) + \sigma_{\text{lag}}^2(\tau, p = 0) = \sigma_{\text{np}}^2(m'_v, \tau) + \sigma_{\text{lag}}^2(\tau, p > 0), \quad (36)$$

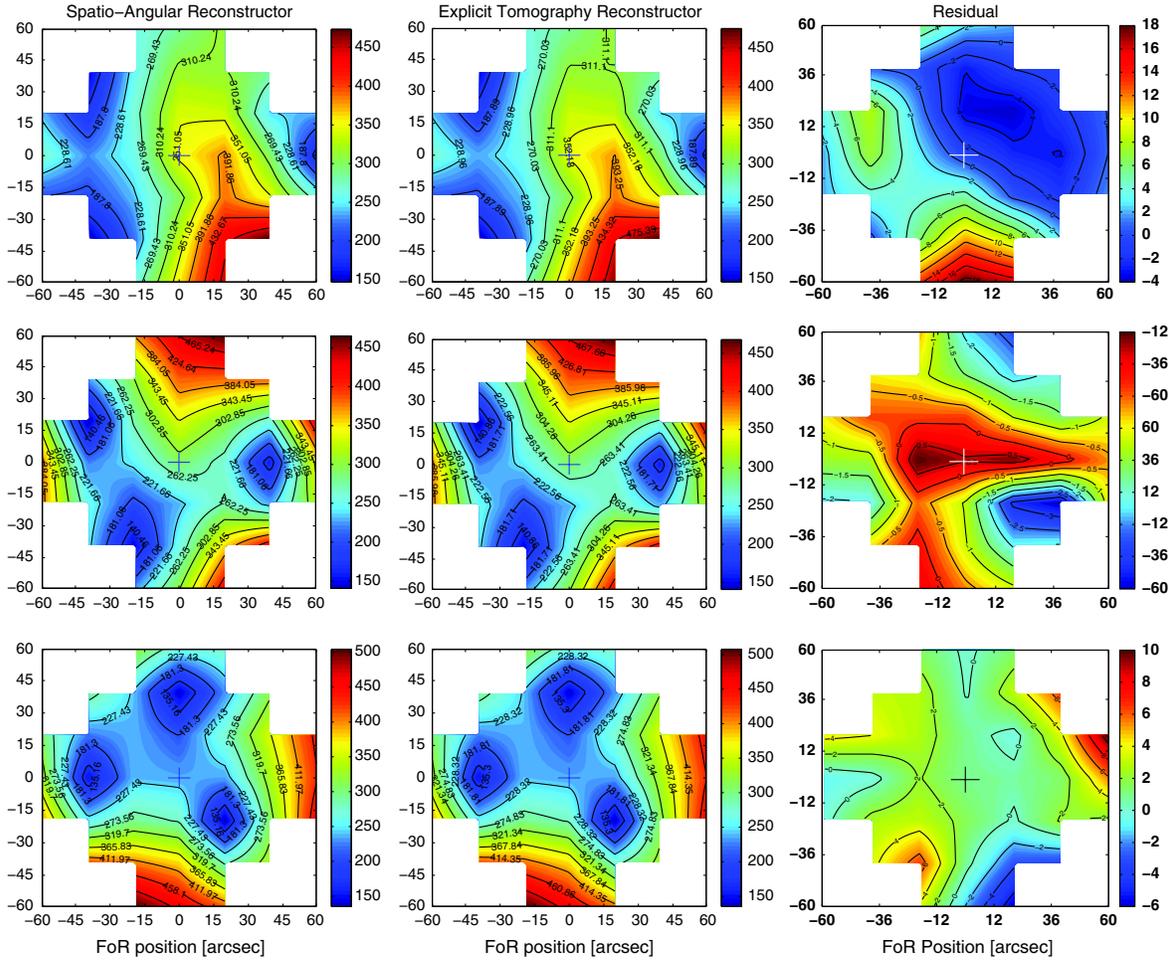


Fig. 7. Comparison of residual wavefront error (in nm RMS) across the FoR with the SA (left column) and the explicit tomographic reconstructor (central column), for NGS asterisms 2, 1.5, and 1 ft apart. Differences plot on the right column.

where  $\sigma_{lag}^2(\tau, p = 0) \geq \sigma_{lag}^2(\tau, p > 0)$  and  $\sigma_{np}^2(m_v, \tau) \leq \sigma_{np}^2(m'_v, \tau)$ , with  $m'_v - m_v$  the limiting magnitude increase.

This rather simplistic approximation offers insight into the potential magnitude increase factor. Figure 8 plots the result,  $m'_v - m_v$ , as a function of lag. The noise-propagation

coefficient used was computed to be  $\sigma_{np}^2/\sigma_n^2 \approx 0.5$ , following an extension to tomography of the analytical derivation in [35].

As expected, the minima in  $\sigma^2$  (in blue, ordinate on the right) are achieved for longer integration times as the GS are fainter. These minima are indicated by vertical dotted

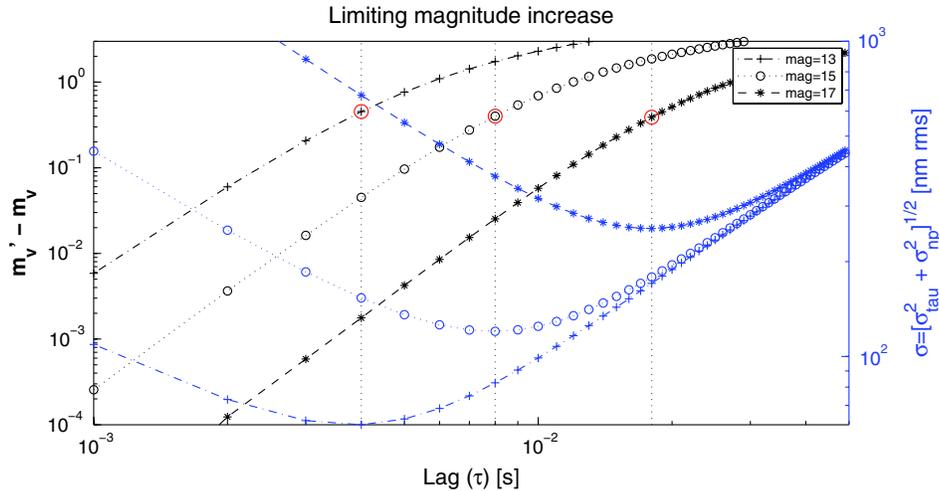


Fig. 8. Temporal lag error is traded by increased noise propagation through the wavefront reconstruction. Black markers: increase in limiting magnitude; blue markers: combined temporal plus noise propagation error; red circles indicate the increased limiting magnitudes for the minima of  $\sigma$ .

**Table 2. Raven End-to-End Simulation Results<sup>a</sup>**

GS Mags	Static MV				Static MV + AR2				Static MV+ One-Step SA				Static MV + AR1			
	EE	Lag	Strehl	Lag	EE	Lag	Strehl	Lag	EE	Lag	Strehl	Lag	EE	Lag	Strehl	Lag
14.0	35.70	7	30.08	7	35.83	9	30.71	10	36.70	9	32.26	10	35.67	8	30.09	8
14.5	35.39	7	28.70	9	35.65	10	29.89	11	36.41	9	30.98	10	35.59	9	28.73	9
15.0	35.32	9	26.69	10	35.38	10	28.38	12	36.00	9	29.74	12	35.11	9	26.71	12
15.5	34.48	10	24.39	12	35.27	12	26.85	14	35.49	12	28.47	17	34.38	10	24.32	13
16.0	33.61	12	21.58	14	34.12	14	24.23	17	35.34	23	26.95	24	33.30	12	21.29	14
16.5	32.24	14	18.05	17	33.59	18	21.60	21	35.10	27	24.81	27	31.98	14	17.97	18
17.0	30.39	18	14.39	27	32.16	21	17.91	28	34.24	33	21.41	33	30.27	19	14.30	28

<sup>a</sup>The optimal performance (% ensquared energy, % Strehl ratio) for each GS magnitude is shown for three reconstructors: the static MV, SA prediction, and the AR2 prediction model. AR1 prediction is included for comparison purposes. Lags quoted in milliseconds.

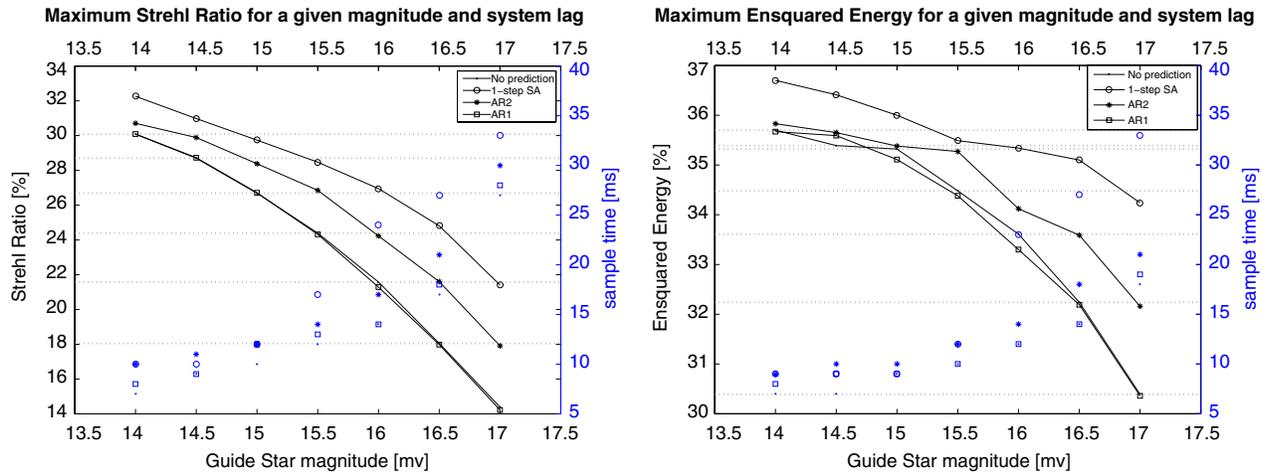


Fig. 9. Left: Strehl ratio. Right: ensquared energy. Peak performance achieved in simulation for GS magnitudes from 14 to 17 using each of four algorithms: static, SA prediction, AR2 prediction, and AR1 prediction (black) and the integration time in ms (ordinates on the right) at which the peak value was reached (blue).

lines, for which red circles indicate the increased limiting magnitudes when the sources vary from magnitudes 13 to 17. A consistent value around half a magnitude increase is obtained. However, only a full end-to-end Monte Carlo simulation can inform about the effective magnitude increase gain, considering the actual EE and Strehl-ratio (SR) figures of merit.

### C. Sample Numerical Simulation

Results of Monte Carlo numerical simulations show the improvement in system performance using EE and SR as the benchmark figures of merit. Simulation parameters were selected to reflect the observing conditions of Raven with the telescope and atmosphere simulator. This includes an asterism of 3 NGSSs within a 2 arcmin FOR; in this case, an asterism with a diameter of 0.5 arcmin was selected in order to reduce tomographic errors and highlight the temporal aspects of system performance. The full set of simulation parameters are listed in Table 1.

The results of the numerical simulations are shown in Fig. 9 and summarized in Table 2. Following Section 4.A, turbulence is expanded on 27 radial orders (i.e., 406 modes) from which nine (55 modes) are controlled. For computational reasons, the predictive models are computed for 90 out of 406 modes per layer with no significant effect on the overall performance.

The data shows the peak system performance in SR and EE, on the Subaru IRCS 140 mas slit, for an asterism of three guide stars of the same magnitude, and the corresponding sampling rate at which the peak occurs for each of three algorithms: static MV, static MV using an AR2 model, and static MV using the one-step SA predictive model. For interest and comparison purposes, the AR1 model also was tested. Lags corresponding to frame rates of 30–200 Hz were used. The simulation incorporates the fixed system lag of 3 ms allotted for camera read-out, data processing, and issuing of DM commands. In order to model the delayed application of the DM commands in the middle of an exposure, the system is attributed a fixed sample rate of 1 kHz. The resulting output phase is summed over the first 3 ms (three samples) before the new DM command is applied; subsequently, the output phase for the remaining total exposure time is added to these first frames to make a total exposure of the desired length. Each simulation run collected 2000 exposures before computing the SR and EE.

The results confirm that a reduction in frame rate, combined with SA temporal prediction (as noted in [11]) will allow the system to achieve a level of performance for a given GS magnitude, which is equal to the performance with a static reconstructor using GSs one magnitude brighter, and the performance with an AR2 predictive reconstructor using GSs half a magnitude brighter. This achievement was underestimated

through the somewhat simplistic model of Eq. (34) used in Fig. 8, for the noise is not spatially invariant, its propagation through the reconstruction isn't constant for varying magnitudes, and neglected terms do indeed depend upon the lag. The results also confirm the computations shown in Fig. 5 that estimate no reduction in temporal error with the use of an AR1 prediction model.

## 5. CONCLUSION AND OUTLOOK

This paper explores the features of two mathematically equivalent versions of algorithms to perform tomography in wide-field AO systems: the commonly used minimum-variance estimators with explicit 3D reconstruction of turbulence profile and the MOAO-specific SA reconstructors adopted in the Learn & Apply from [4] where the wavefront is directly estimated in the pupil plane.

A merge of both formulations is proposed to overcome (1) a challenging calibration since MOAO systems operate in OL and (2) intrinsic temporal lag errors by embedding a predictive model to work with fainter sources and thus increase the sky coverage. The former is tackled by partially computing the tomographic reconstructor from OL measurements. This mitigates issues related to under modeling and unknown systematics between the WFSs. The latter needs always to be model-based, with parameters identified by postprocessing the measurements' covariance matrix with the built-in SLODAR method. This renders the tomographic reconstruction a highly data-driven, self-sufficient approach to circumvent the complex calibration in OL operation.

Several predictive models are described along with a much faster and computationally sound identification procedure over previous methods (assuming Zernike modes). The predicted error structure functions are analytically derived. Diagonal AR models of order 2 are shown to perform equally well as the SA predictor model for small lags below  $\approx 10$  ms. For larger lags, the one-step SA predictor is preferred.

These models can now be plugged into the linear-quadratic-Gaussian controller, which is expected to push further the limiting magnitude. At this stage, the wind profiles are assumed known, and only pure frozen flow is simulated. In reality, the profiles are estimated and frozen flow may not be always present, which calls for a robustness assessment to be done in a forthcoming paper.

For instance, on the Raven project, results in this paper suggest the same level of performance is achieved by combining temporal prediction with guide stars 1 magnitude fainter.

## ACKNOWLEDGMENTS

The authors acknowledge support by the European Research Council/European Community under the FP7 through starting grant agreement no. 239953 and the Marie Curie Intra-European Fellowship grant with reference FP7-PEOPLE-2011-IEF, no. 300162. Thanks to Rodolphe Conan for many discussions and for providing and maintaining the object-oriented MATLAB-based AO simulator (OOMAO) freely available from <https://github.com/rconan/OOMAO/>

## REFERENCES

1. D. R. Andersen, K. J. Jackson, C. Blain, C. Bradley, C. Correia, M. Ito, O. Lardi re, and J.-P. V eran, "Performance modeling for the raven multi-object adaptive optics demonstrator," *Publ. Astron. Soc. Pac.* **124**, 469–484 (2012).
2. F. J. Rigaut, B. L. Ellerbroek, and R. Flicker, "Principles, limitations, and performance of multiconjugate adaptive optics," *Proc. SPIE* **4007**, 1022–1031 (2000).
3. A. Tokovinin, M. Le Louarn, E. Viard, N. Hubin, and R. Conan, "Optimized modal tomography in adaptive optics," *Astron. Astrophys.* **378**, 710–721 (2001).
4. F. Vidal, E. Gendron, and G. Rousset, "Tomography approach for multi-object adaptive optics," *J. Opt. Soc. Am. A* **27**, A253–A264 (2010).
5. M. R. Whiteley, B. M. Welsh, and M. C. Roggemann, "Optimal modal wave-front compensation for anisoplanatism in adaptive optics," *J. Opt. Soc. Am. A* **15**, 2097–2106 (1998).
6. G. C. Valley, "Long- and short-term Strehl ratios for turbulence with finite inner and outer scales," *Appl. Opt.* **18**, 984–987 (1979).
7. B. L. Ellerbroek, "Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques," *J. Opt. Soc. Am. A* **19**, 1803–1816 (2002).
8. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am. A* **66**, 207–211 (1976).
9. R. Ragazzoni, E. Marchetti, and F. Rigaut, "Modal tomography for adaptive optics," *Astron. Astrophys.* **342**, L53–L56 (1999).
10. T. Fusco, J.-M. Conan, G. Rousset, L. M. Mugnier, and V. Michau, "Optimal wave-front reconstruction strategies for multiconjugate adaptive optics," *J. Opt. Soc. Am. A* **18**, 2527–2538 (2001).
11. P. Piatrou and M. C. Roggemann, "Performance study of Kalman filter controller for multiconjugate adaptive optics," *Appl. Opt.* **46**, 1446–1455 (2007).
12. L. Gilles, "Closed-loop stability and performance analysis of least-squares and minimum-variance control algorithms for multiconjugate adaptive optics," *Appl. Opt.* **44**, 993–1002 (2005).
13. B. D. O. Anderson and J. B. Moore, *Optimal Filtering* (Dover, 1995).
14. G. C. Valley and S. M. Wandzura, "Spatial correlation of phase-expansion coefficients for propagation through atmospheric turbulence," *J. Opt. Soc. Am.* **69**, 712–717 (1979).
15. F. Chassat, "Calcul du domaine d'isoplan tisme d'un syst me d'optique adaptative fonctionnant   travers la turbulence atmosph rique," *J. Opt.* **20**, 13–23 (1989).
16. N. Takato and I. Yamaguchi, "Spatial correlation of Zernike phase-expansion coefficients for atmospheric turbulence with finite outer scale," *J. Opt. Soc. Am. A* **12**, 958–963 (1995).
17. 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," *J. Opt. Soc. Am. A* **15**, 993–1005 (1998).
18. D. M. Winker, "Effect of a finite outer scale on the Zernike decomposition of atmospheric optical turbulence," *J. Opt. Soc. Am. A* **8**, 1568–1573 (1991).
19. A. Cort s, B. Neichel, A. Guesalaga, J. Osborn, F. Rigaut, and D. Guzman, "Atmospheric turbulence profiling using multiple laser star wavefront sensors," *Mon. Not. R. Astron. Soc.* **427**, 2089–2099 (2012).
20. V. P. Lukin, "Dynamics of adaptive optical systems," *J. Opt. Soc. Am. A* **27**, A216–A222 (2010).
21. R. Ragazzoni, E. Diolaiti, J. Farinato, E. Fedrigo, E. Marchetti, M. Tordi, and D. Kirkman, "Multiple field of view layer-oriented adaptive optics," *Astron. Astrophys.* **396**, 731–744 (2002).
22. F. Ass mat, R. Wilson, and E. Gendron, "Method for simulating infinitely long and non stationary phase screens with optimized memory storage," *Opt. Express* **14**, 988–999 (2006).
23. D. L. Fried and T. Clark, "Extruding kolmogorov-type phase screen ribbons," *J. Opt. Soc. Am. A* **25**, 463–468 (2008).
24. A. Beghi, A. Cenedese, and A. Masiero, "Stochastic realization approach to the efficient simulation of phase screens," *J. Opt. Soc. Am. A* **25**, 515–525 (2008).
25. C. Kulcs r, H.-F. Raynaud, C. Petit, J.-M. Conan, and P. V. de Lesegno, "Optimal control, observers and integrators in adaptive optics," *Opt. Express* **14**, 7464–7476 (2006).
26. G. Sivo, H.-F. Raynaud, J.-M. Conan, C. Kulcsr, E. Gendron, F. Vidal, and A. Basden, "First laboratory validation of LQG control with the CANARY MOAO pathfinder," *Proc. SPIE* **8447**, 84472Y (2012).

27. C. Correia, J.-P. Véran, G. Herriot, B. L. Ellerbroek, L. Wang, and L. Gilles, "Increased sky coverage with optimal correction of tilt and tilt-anisoplanatism modes in laser-guide-star multiconjugate adaptive optics," *J. Opt. Soc. Am. A* **30**, 604–615 (2013).
28. C. Correia, H.-F. Raynaud, C. Kulcsár, and J.-M. Conan, "On the optimal reconstruction and control of adaptive optical systems with mirror dynamics," *J. Opt. Soc. Am. A* **27**, 333–349 (2010).
29. B. Le Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. M. Mugnier, and T. Fusco, "Optimal control law for classical and multiconjugate adaptive optics," *J. Opt. Soc. Am. A* **21**, 1261–1276 (2004).
30. J.-M. Conan, G. Rousset, and P.-Y. Madec, "Wave-front temporal spectra in high-resolution imaging through turbulence," *J. Opt. Soc. Am. A* **12**, 1559–1570 (1995).
31. K. Jackson, C. Correia, O. Lardière, D. Andersen, and C. Bradley, "Tomography for Raven, a multi-object adaptive optics science and technology demonstrator," in *Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS 2012)*, Maui, Hawaii, September 11–14, 2012 (Curran Associates, 2012).
32. C. Petit, J.-M. Conan, C. Kulcsár, and H.-F. Raynaud, "Linear quadratic Gaussian control for adaptive optics and multiconjugate adaptive optics: experimental and numerical analysis," *J. Opt. Soc. Am. A* **26**, 1307–1325 (2009).
33. A. C. Robin, C. Reylé, S. Derrière, and S. Picaud, "A synthetic view on structure and evolution of the Milky Way," *Astron. Astrophys.* **409**, 523–540 (2003).
34. R. M. Clare, B. L. Ellerbroek, G. Herriot, and J.-P. Véran, "Adaptive optics sky coverage modeling for extremely large telescopes," *Appl. Opt.* **45**, 8964–8978 (2006).
35. F. Rigaut and E. Gendron, "Laser guide star in adaptive optics—the tilt determination problem," *Astron. Astrophys.* **261**, 677–684 (1992).