



CLOSED-LOOP ADAPTIVE OPTICS CONTROL IN
STRONG ATMOSPHERIC TURBULENCE

DISSERTATION

Todd M. Venema, Lieutenant Colonel, USAF

AFIT/DEE/ENG/08-21

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/DEE/ENG/08-21

CLOSED-LOOP ADAPTIVE OPTICS CONTROL IN
STRONG ATMOSPHERIC TURBULENCE

DISSERTATION

Presented to the Faculty
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

Todd M. Venema, B.S.E., M.S.E.E.
Lieutenant Colonel, USAF

September 2008

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

CLOSED-LOOP ADAPTIVE OPTICS CONTROL IN
STRONG ATMOSPHERIC TURBULENCE

Todd M. Venema, B.S.E., M.S.E.E.
Lieutenant Colonel, USAF

Approved:

_____	_____
Dr. Juan Vasquez, (Chairman)	Date

_____	_____
Maj. Jason Schmidt, PhD (Member)	Date

_____	_____
Dr. Matthew Fickus (Member)	Date

Accepted:

_____	_____
M. U. Thomas	Date
Dean, Graduate School of	
Engineering and Management	

Abstract

A self-referencing interferometer based closed-loop adaptive optics controller is developed which is designed to operate effectively under strong turbulence conditions. The aberrated optical field is modeled stochastically and then estimates of the state of the system are developed using a steady-state, fixed-gain Kalman filter. The phase of the optical field is considered the state of the system which is wrapped in a limited range of $(-\pi, \pi]$. This phase is unwrapped through the use of a least-squares reconstructor which has been modified to work effectively in the presence of branch points associated with strong turbulence. The conjugate of the optical phase is then applied to the system's deformable mirror in order to correct for the effects of atmospheric turbulence on the optical field.

The advances developed in this research are in the application of a steady-state, fixed-gain Kalman filter to the input of an adaptive optic system, unwrapping the optical phases after the field estimation, and improving the phase unwrapping by varying the domain of the rotational phase component present in strong turbulence.

The system developed in this research is shown in computer simulation to be improved over current designs by comparing performance plots of system Strehl ratios for systems utilizing the different designs.

For my mom and dad who led the way, for my wife who walks with me, for my kids who follow. Thank you Lord for lighting the path.

Acknowledgements

I would like to acknowledge the help of Jason Schmidt and Juan Vasquez, my Air Force Institute of Technology advisors. I would also like to acknowledge the help of Darryl Sanchez and Denis Oesch from the Air Force's Starfire Optical Range in helping me study my designs in their Atmospheric Simulation and Adaptive-optics Laboratory Testbed (ASALT).

Todd M. Venema

Table of Contents

	Page
Abstract	iv
Acknowledgements	vi
Table of Contents	vii
List of Figures	x
List of Tables	xv
 I. Introduction	 1
1.1 AO systems	1
1.2 Shack-Hartmann Wavefront Sensor	2
1.3 Self-Referencing Interferometers	3
1.4 Motivation	6
 II. Literature Review	 7
2.1 Adaptive Optics	7
2.2 Wavefront Sensors	7
2.2.1 Shack-Hartmann Wavefront Sensor	8
2.2.2 Temporally Phase-Shifted Self-Referencing Interferometer	10
2.2.3 Spatial SRI	15
2.2.4 SRI performance	16
2.3 Wavefront Correction Devices	17
2.3.1 Continuous Facesheet Deformable Mirrors	17
2.3.2 Segmented Deformable Mirrors	17
2.4 Stochastic Control	18
2.4.1 State Space System Representation	19
2.4.2 Stochastic Modeling	19
2.4.3 Stochastic Estimation	23
2.4.4 Stochastic Control	24
2.4.5 LQG	24
2.5 Continuity of the Optical Field	26
2.6 Phase Discontinuities	27
2.7 Unwrapping Phase Discontinuities	28
2.8 Branch Points	28
2.9 Rotational and Irrotational Field Components	30
2.10 Difficulties of Branch Points	30
2.10.1 Wavefront Reconstruction	30
2.10.2 DMs and the irrotational phase	30
2.11 Branch Cuts	30

	Page
2.11.1 Branch cut impact on DMs	31
2.12 Spatial sampling requirements	31
2.13 SRI Development	32
2.14 Alternative viewpoint of phase discontinuities	32
2.15 Chapter conclusions	34
III. Adaptive optics, the optical Kalman filter	46
3.1 Introduction	46
3.2 Estimation and correction versus control	46
3.3 Modeling the system	52
3.3.1 System states	52
3.3.2 System Dynamics	53
3.3.3 Measurement noise	54
3.4 Deriving wavefront amplitude from SRI output	64
3.5 Kalman Filter	70
3.5.1 Linear Kalman Filter or Extended Kalman Filter	70
3.6 Chapter Conclusions	71
IV. When to Unwrap	72
4.1 Introduction	72
4.2 Weak Versus Strong Turbulence AO Systems	72
4.2.1 Weak-turbulence AO systems	72
4.2.2 Strong-turbulence AO system	73
4.2.3 Uncharted islands	74
4.2.4 Island persistence	74
4.2.5 Dealing with the problem	75
4.3 Chapter conclusions	78
V. Optical phase unwrapping in the presence of branch points	88
5.1 Introduction	88
5.1.1 Phase Cuts	88
5.1.2 Wrapping Cuts	88
5.1.3 Branch Cuts	89
5.1.4 Least-Squares Unwrappers	90
5.1.5 Non-LS Component of the Field	92
5.2 Improved Unwrapper	93
5.2.1 Unwrapping Metric - Normalized Cut Length	95
5.3 Simulation and Results	97
5.4 Comparison to Other Unwrappers	100
5.5 Impact on System Performance	101
5.6 Chapter Conclusion	102

	Page
VI. Results	106
6.1 Introduction	106
6.2 Basic AO structure	106
6.3 Varying r_0	106
6.4 Varying sample rates	108
6.5 Varying read noise	110
6.6 Chapter conclusions	113
VII. Conclusions	115
7.1 Significant Contributions	115
7.1.1 Kalman estimation of anisoplanatic Zernike tilt	115
7.1.2 An improved temporally phase-shifted design	115
7.1.3 Recognition of the AO controller as an estimator	115
7.1.4 Unwrapping last	115
7.1.5 Improved unwrapping	116
7.2 A single graph	116
7.3 Future Work	117
Appendix A. Appendix	119
A.1 Real and Imaginary contour generator code	119
A.2 Branch point finder	124
A.3 pdf maker	125
A.4 Error variance for different phases	126
A.5 Variance generator	128
Bibliography	132

List of Figures

Figure		Page
1.1.	Example of a simple open-loop AO system	2
1.2.	Open-loop control system block diagram. ‘In’ represents the light incident on the telescope while ‘Out’ represents the light incident on the imaging camera.	2
1.3.	Conventional AO System [40]	3
1.4.	Closed-Loop control System	3
1.5.	Uncorrected image (left) vs. AO corrected image (right) [45]	4
1.6.	Performance comparison between S-H WFS with conventional Least Squares reconstructor (green line), SRI WFS with exponential filter then unwrapper (blue line and top left DM depiction), and SRI WFS with unwrapping the linear filtering (red line and top right DM depiction) [34].	5
7.	Shack-Hartmann lenslet diagram [40]	9
8.	Determining phase tilt from a S-H WFS [8]	10
9.	Conceptual diagram of a temporally phase-shifted SRI [35]	11
10.	Conceptual Spatially Designed SRI [37]	15
11.	Frame sequence for a temporal SRI [38]. The black outline indicates the extent of the camera. The blue circle is the area of the pupil interferogram.	16
12.	Spatial SRI [34]	16
13.	Continuous Facesheet DM [40]	17
14.	Segmented DM [40]	18
15.	State space system representation [28]	19
16.	Kalman filter illustration [30]	22
17.	LQG controller illustration [30]	25

Figure		Page
18.	Four quadrant $\tan^{-1}(x, y)$ function. The closed circle at π indicates that $\tan^{-1}(0, x) = \pi \forall x < 0$ while the open circle just below indicates that $\lim_{y \rightarrow 0^-} \tan^{-1}(y, x) = -\pi \forall x < 0$. Thus, the function is discontinuous at $y = 0, x \leq 0$, since $\lim_{y \rightarrow 0^-} \tan^{-1}(y, x) \neq \tan^{-1}(0, x) \forall x < 0$ and $\tan^{-1}(0, 0)$ is undefined.	27
19.	Phases for a single row of a wrapped optical field. The phases are restricted to $(-\pi, \pi]$ from the MATLAB angle function which can be expressed as $\text{angle}(z) = \text{atan2}(\text{imag}(z), \text{real}(z))$ where z is the complex variable given to the function.	28
20.	Phases for a single row of an unwrapped optical field. The phases are no longer restricted to $(-\pi, \pi]$. Integer multiples of 2π are added to the phase in order to smooth the phase in a process known as unwrapping.	29
21.	An unwrapping process. First a single column is unwrapped in the center of the field. This avoids spurious data in the corners of the field due to the circular mask of the aperture. Next the field is unwrapped from the center column to the outsides. The right half is unwrapped first, then the array is flipped and the remaining half is unwrapped.	36
22.	Phase gradients for a field containing branch points.	37
23.	Phase gradients for a field without branch points.	38
24.	A slice through the real portion of an optical field where the real portion of the field equals zero. The optical field is a 50×50 section of an optical field generated by Wavetrain [®] of an idealized point source through weak atmospheric turbulence.	39
25.	A slice through the imaginary portion of an optical field where the imaginary portion of the field equals zero. The optical field is the same 50×50 section of an optical field depicted by Figure 24.	40
26.	An overlay plot of the real and imaginary contours slices in Figures 24 and 25. Note the areas where two similar contours separated by an opposite contour (blue-red-blue) are only two pixels apart. This indicates that the field is at best critically sampled.	41
27.	A sliced contour plot for a field corrupted by atmospheric turbulence with Rytov number 0.1. Note the increased spacing between contour lines as compare with Figure 26 indicating a more comfortable spatial sampling rate.	42

Figure		Page
28.	A sliced contour plot for a field corrupted by atmospheric turbulence with Rytov number 0.4. Note the presence of branch points in the field.	43
29.	Frame 1 - A depiction of real and imaginary contour slices and the branch point produced where they intersect.	44
30.	Frame 2 - Note that the branch points are still paired similar to frame 1 in Figure 29, but they are moving closer together as the circles separate.	44
31.	Frame 3 - The contours have moved to the point where the lines of the circles no longer cross but only touch. Thus the circles do not cause a branch point, but new branch points are created by the real circle and an imaginary portion = 0 line that spans the aperture from edge to edge.	45
32.	Block diagram of a leaky integrator	47
33.	Stochastic model of turbulence with DM	47
34.	Optimal control structure block diagram	48
35.	Optimal control structure	48
36.	Portion of optimal control structure	49
37.	Controlled and Corrected systems	50
38.	Block diagram of AO estimator	50
39.	Revised system model	51
40.	Revised Kalman filter	51
41.	Final system design	52
42.	Histogram of subaperture phase changes from frame to frame	54
43.	Graphical depiction of SRI phase estimation	55
44.	Empirical pdf of ϕ_{SRI} with moderate noise	56
45.	Empirical pdf of ϕ_{SRI} with high noise	57
46.	Standard deviation σ vs ϕ	58
47.	Variance of ϕ_{SRI} determined through Monte Carlo analysis	59
48.	Variance of ϕ_{SRI} determined through Equation (32)	60
49.	K vs. $\frac{R}{Q_d}$	63

Figure		Page
50.	K vs. A_R and A_S	64
51.	Graphical depiction of SRI phase estimation	65
52.	Normalized reference intensity	67
53.	Threshold energy and total energy	68
54.	Subaperture segregation: Red = $A_R > A_S$, Blue = $A_R < A_S$	69
55.	Block diagram of a S-H based AO system	73
56.	Example residual phase showing 2π islands. Lines of blue/red indicate where the DM is making 2π transitions. Note the isolated circle just slightly above center. The circular shape and the double transition (blue-red-green-blue-red-green) indicate that a single DM actuator is displaced 4π from what it should be.	75
57.	Strehl ratio performance of ‘unwrap then control’ design	76
58.	Buildup of phase cuts in ‘unwrap then control’ AO	80
59.	Residuals after 500 frames for varying levels of A	81
60.	Comparison of various leak levels on system performance	82
61.	Block diagram of ASALT lab setup	82
62.	Screen shot of ASALT lab graphical user interface	83
63.	Interferogram showing flat DM (open-loop)	84
64.	Interferogram of DM after 100 frames in AO system with A = 1.0	84
65.	Interferogram of DM after 100 frames in AO system with A = 0.998	85
66.	Interferogram of DM after 100 frames in AO system with A = 0.99	85
67.	Interferogram of DM after 100 frames in AO system with A = 0.95	86
68.	Interferogram of DM after 100 frames in AO system with A = 0.9	86
69.	Block diagram of a SRI based ‘Control then Unwrap’ AO system	87
70.	Performance of ‘Control then unwrap’ design	87
71.	(a) Wrapped phase with only wrapping cuts. (b) Unwrapped version of (a). Note that the unwrapped phase is smooth.	89
72.	Wrapped phase with both wrapping and branch cuts. If this phase were to be unwrapped, it would not be smooth.	90

Figure		Page
73.	Poor unwrap and minimum cut distance unwrap of phase field with branch points.	91
74.	Intensity overlaid by branch cuts using LS unwrapper to eliminate wrapping cuts.	93
75.	Poor unwrapping, phase and intensity overlaid by branch cuts.	94
76.	Branch cuts of four different unwrap realizations.	96
77.	Field estimation Strehl ratio versus integrated cut intensity.	97
78.	Comparison of closed-loop AO performance between variable and fixed ϕ_{non-LS} range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrappers.	104
79.	CDF comparisons between variable range ϕ_{non-LS} and fixed range ϕ_{non-LS} unwrappers.	105
80.	Block diagram of the AO system.	106
81.	Depiction of WFS subaperture and DM actuator positions.	107
82.	System performance (Strehl ratio) for $D_{SA}/r_0 = 0.25$	108
83.	System performance (Strehl Ratio) vs. D_{SA}/r_0	109
84.	System performance (Strehl Ratio) vs. normalized sampling rate	110
85.	Effect of varying the Kalman gain K for low noise, low sample rate	111
86.	Interferogram intensity with colorbar scale	112
87.	Interferogram noise variance with colorbar scale	113
88.	System performance versus measurement noise	114
89.	System performance at high noise levels depicting slow lock onto signal	114
90.	Comparison between AO systems	117

List of Tables

Table		Page
1.	Normalized cut lengths for 0.4 log-amplitude variance field	99
2.	Normalized cut lengths for 0.8 log-amplitude variance field	99
3.	Normalized cut lengths from various unwrappers, 0.4 log-amplitude variance field	101
4.	Normalized cut lengths from various unwrappers, 0.8 log-amplitude variance field	101
5.	Average Strehl results for 1000 frame simulation	102

CLOSED-LOOP ADAPTIVE OPTICS CONTROL IN STRONG ATMOSPHERIC TURBULENCE

I. Introduction

Adaptive Optics (AO) is used to correct for the effect of atmospheric turbulence on an optical system. By correcting for the effect of atmospheric turbulence, the system can be improved until it becomes diffraction-limited, at which point it performs as if looking through a vacuum instead of turbulent air. In this case, the resolution of the system is limited solely by the aperture size of the system.

1.1 AO systems

AO systems have three main parts. First, a wavefront sensor (WFS) measures the wavefront of light received by the system. Second, a controller takes the output from the WFS and creates the input to the third part of an AO system, a wavefront compensator which corrects the wavefront. [12]

Figure 1.1 shows a simple system in which the distorted wavefront is sensed by a WFS. Wavefront measurements are given to the controller which then commands a deformable mirror (DM) to minimize the effect of atmospheric turbulence. This is an open-loop system which is conceptually valuable but not very effective in practice. A block diagram of the equivalent control loop is depicted in Figure 1.2.

A more realistic depiction of an adaptive optics system is given in Figure 1.3, with the equivalent control loop block diagram in Figure 1.4. This system is closed-loop because the DM corrects the wavefront prior to it encountering the WFS. In this system a non-deformable mirror known as a fast-steering mirror (FSM) corrects the tilt (average phase gradient over the entire aperture). A second mirror is deformable and flattens the wavefront. This arrangement minimizes the dynamic range required of the DM.

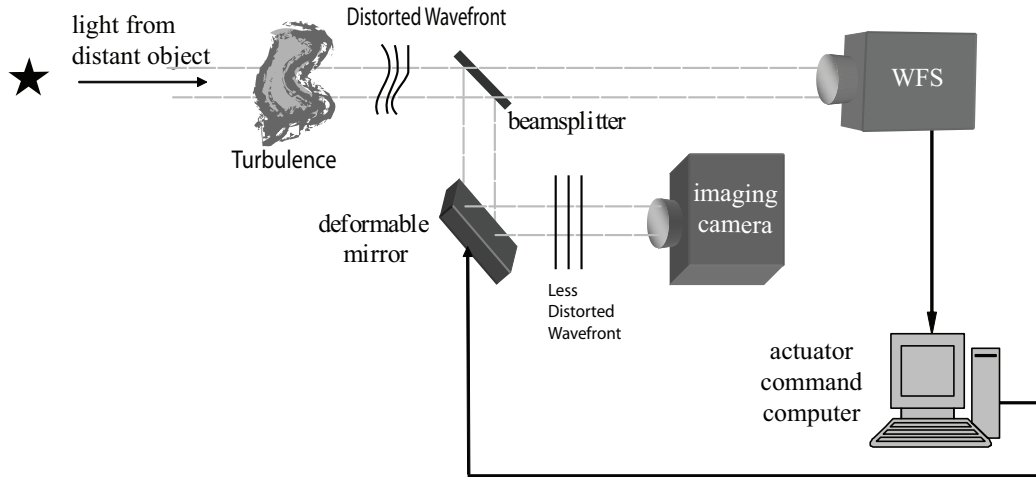


Figure 1.1: Example of a simple open-loop AO system

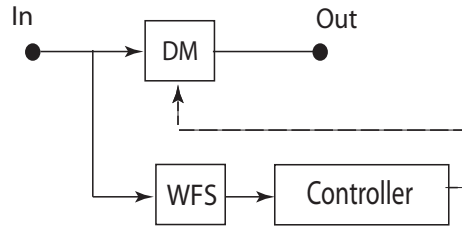


Figure 1.2: Open-loop control system block diagram. ‘In’ represents the light incident on the telescope while ‘Out’ represents the light incident on the imaging camera.

1.2 Shack-Hartmann Wavefront Sensor

Most AO systems use a Shack-Hartmann (S-H) WFS (which measures the field gradient of light), a controller and a DM. Correspondingly, improvements to AO systems have been achieved by improving either the S-H WFS, DM, or the control algorithm connecting the two. After years of development, systems of this design are relatively mature and offer good performance in correcting for weak atmospheric turbulence.

Figure 1.5 shows two images of a binary star system taken at AFRL’s Starfire Optical Range (SOR). “These images were taken using the SOR’s 3.5 m telescope in the I band, which has a center wavelength of ~ 850 nm. The compensated image was corrected by the 756 active actuator AO system. The angular separation between the two stars is 1.45 mrad. The top two images are auto scaled, but the surface plots on the bottom are on the same

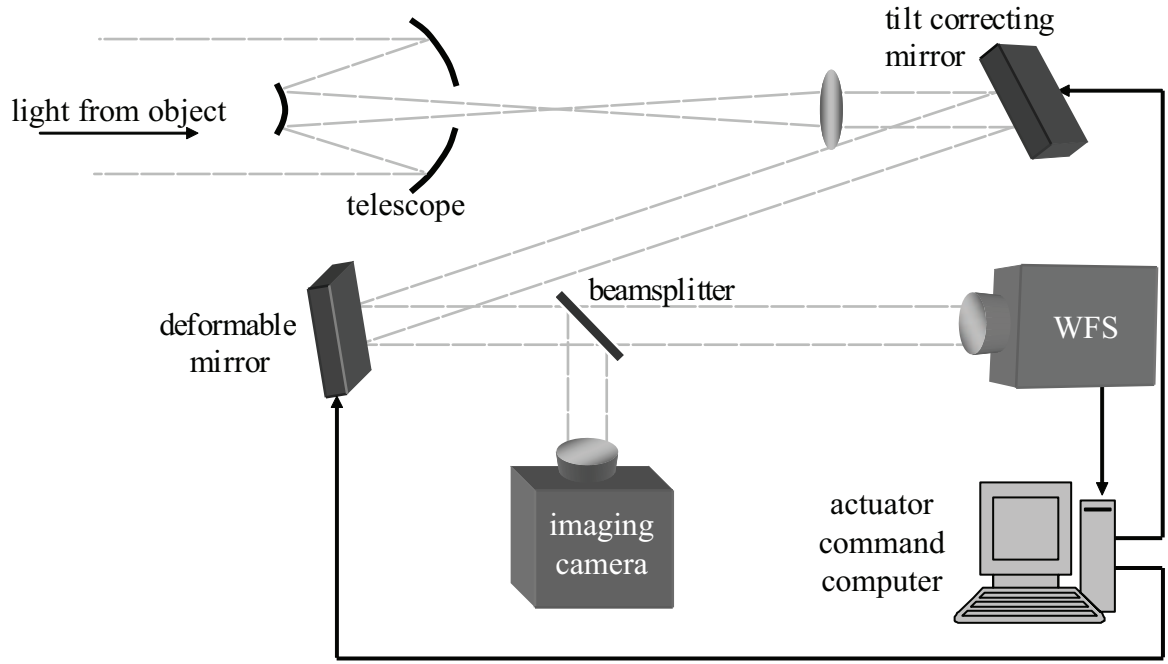


Figure 1.3: Conventional AO System [40]

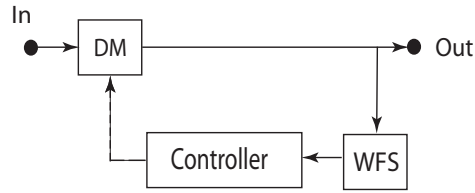


Figure 1.4: Closed-Loop control System

scale. Notice that the peak intensity is much, much greater in the compensated image.” [45] This clearly illustrates how effective conventional AO can be in weak turbulence.

1.3 Self-Referencing Interferometers

Recently, a new type of WFS known as a self-referencing interferometer (SRI) has been developed. An SRI is an appealing alternative to the S-H because it more directly senses the optical field while the S-H simply measures the field gradient. An SRI has two distinct advantages over a S-H WFS. The primary benefit of an SRI is better performance under strong turbulence conditions. [24] Traditional WFSs and reconstruction methods ignore a portion of the phase caused by strong turbulence (the rotational field, explained in Chapter II) while an SRI measures it accurately. [24] A second benefit to an SRI is that by more

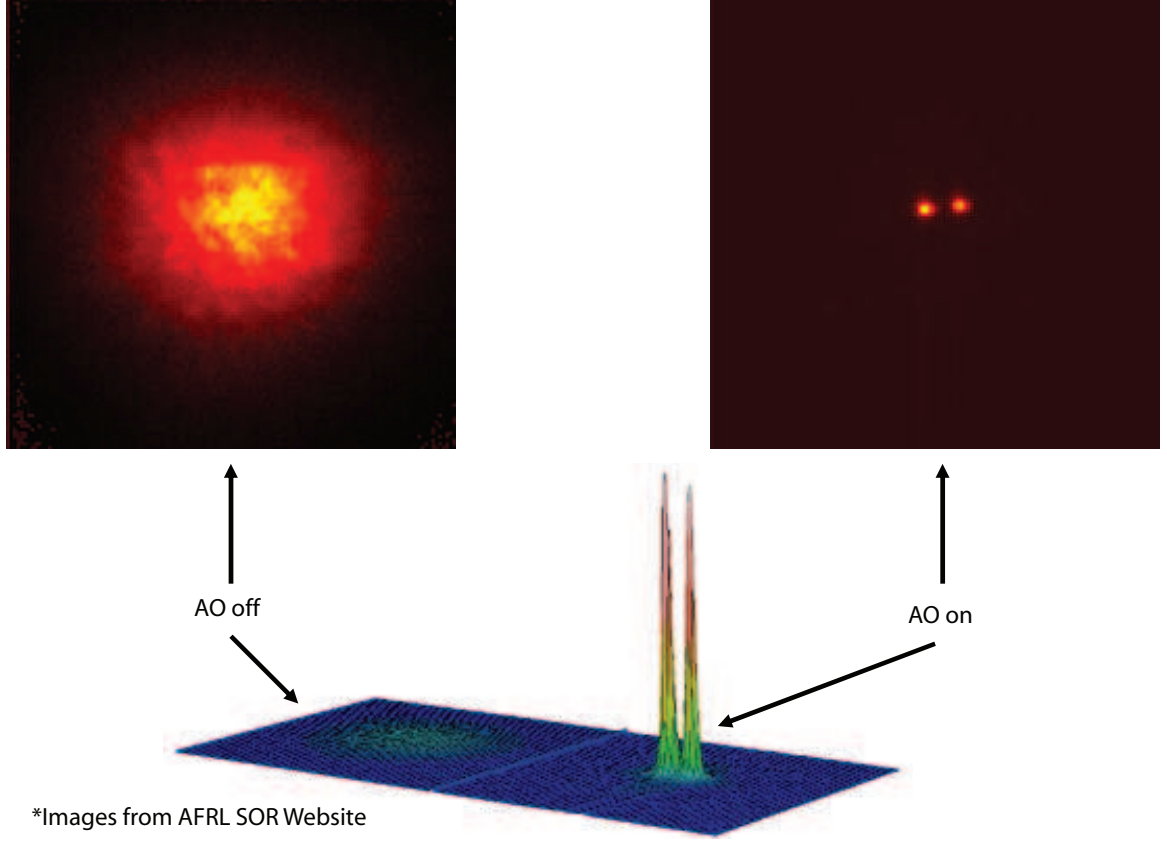


Figure 1.5: Uncorrected image (left) vs. AO corrected image (right) [45]

directly sensing the field, the reconstruction of the field from the WFS output is greatly simplified.

For weak turbulence, the two types of WFSs provide comparable performance. The S-H WFS gradient measurements can be reconstructed into phases with relative ease and the turbulence is weak enough that the traditional systems detect the entire field. At stronger turbulence levels, however, effective AO control becomes much more difficult for traditional systems.

In strong turbulence conditions, the phase of the received light has significant spatial variation. In addition, a phenomenon known as scintillation starts to occur where the amplitude of the optical field varies causing bright spots and nulls to appear in the field. These nulls can lead to something called branch points (explained in more detail in Chapter II) where the phase of the optical field becomes discontinuous. This complicates all three as-

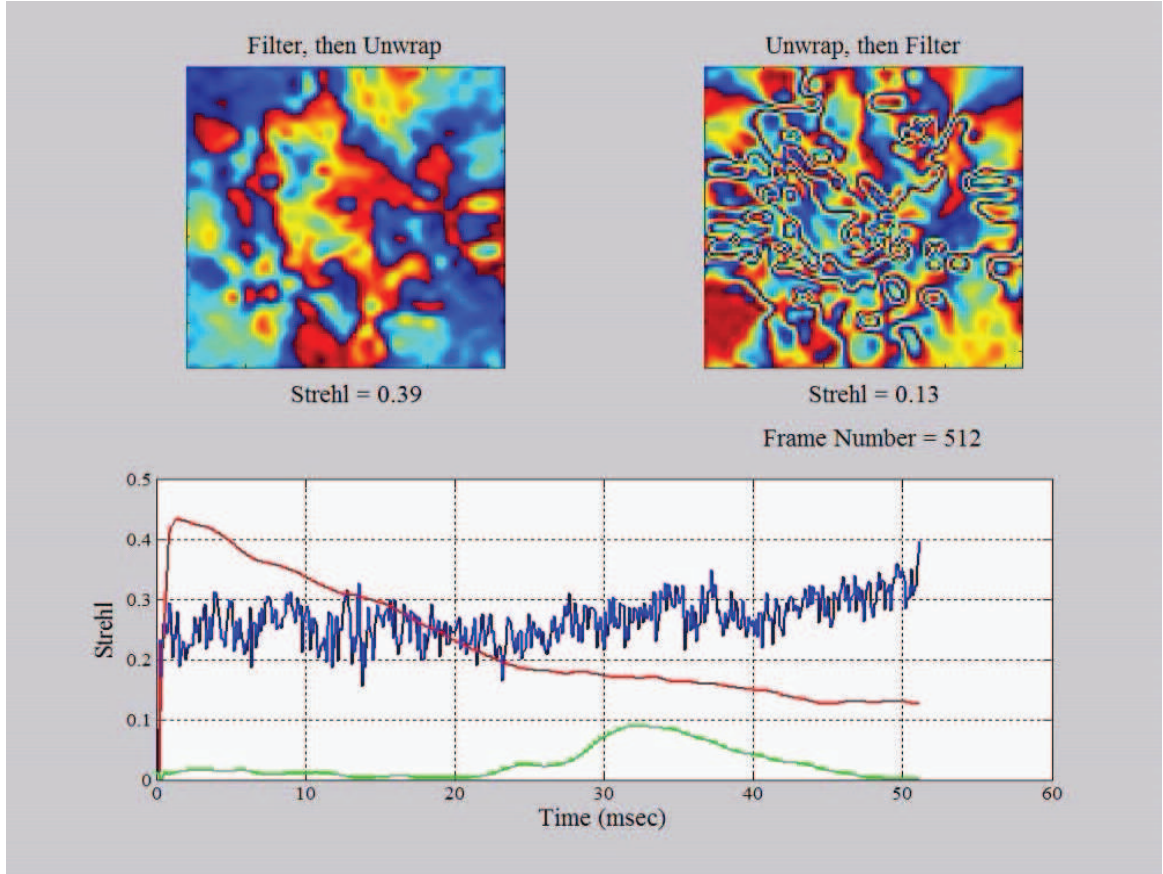


Figure 1.6: Performance comparison between S-H WFS with conventional Least Squares reconstructor (green line), SRI WFS with exponential filter then unwrapper (blue line and top left DM depiction), and SRI WFS with unwrapping the linear filtering (red line and top right DM depiction) [34].

pects of standard AO systems: the reconstruction of the field gradient measurements into the field, the computation of optimum control of the field, and the control of the field.

Since an SRI effectively eliminates the need for wavefront reconstruction and senses the rotational component of the phase, it promises a significant improvement over gradient sensors for these stronger turbulence conditions. However, using the new sensor in closed-loop control has proven to be problematic. Ignoring the branch point effects limits performance of the system. Including branch point effects has created stability problems in the control loops.

The bottom plot in Figure 1.6 show AO performance for three different system designs under similar strong turbulence atmospheric conditions. The green line shows a system using

a S-H WFS and a least-squares (LS) reconstructor. The poor performance of this system is due to ignoring the effect of branch points on the measured phase of the field. The blue line depicts the performance of system using an SRI with an exponential filter and an unwrapper designed to account for branch point effects. This system takes branch point effects into account, which causes it to significantly outperform the first system. The variability in the performance of the system, however, is due to the system constructing differing branch cuts from frame to frame. The correction present on the DM at the end of the test is depicted on the top left of the figure. The red line shows performance of a system utilizing an SRI with an unwrapper and a linear filter. Branch points are taken into account, but the system becomes unstable and eventually underperforms the S-H WFS and exponential filter system. This instability is thought to be due to the undersampling of the optical field by the SRI which leads to erroneously identifying branch points which do not in fact exist. The erroneous branch points build up on the DM because they cannot be sensed and eventually yield a de-stabilized DM like the one depicted on the top right of the figure. [25]

The goal of this dissertation then is to develop a closed-loop AO control structure effective under strong turbulence conditions. The design utilizes an SRI WFS with a controller which precludes buildup of unnecessary branch points effects on the DM.

1.4 Motivation

Weapons are the tools of the warrior and just as pilots need day/night all-weather aircraft, military systems such as the Airborne Laser (ABL) which use optically directed energy must work in strong turbulence conditions. A well-designed SRI-based AO system should perform effectively under strong turbulence conditions, while conventional AO systems cannot.

What if free people could live secure in the knowledge that their security did not rest upon the threat of instant U.S. retaliation to deter a Soviet attack, that we could intercept and destroy strategic ballistic missiles before they reached our own soil or that of our allies? - Ronald Reagan [33]

II. Literature Review

This chapter covers the background material and relevant research literature for the project. It covers all the parts of a closed-loop AO system with special emphasis on the WFS. More specifically, S-H and SRI WFSs are covered in detail. Stochastic estimation and control is covered, and the workings of a Kalman filter are described. Wavefront manipulation devices are discussed. Finally, branch point phenomena and their effects on AO controllers are covered.

2.1 Adaptive Optics

The effects of atmospheric turbulence on optical systems were noted as early as 1656 by Christian Huygens [21]. Automatic tracking of objects, however, was not accomplished until the 1950s and compensating for atmospheric turbulence was first suggested in 1953 by Horace W. Babcock [21]. Later, in 1973, the first crude AO system was tested at the Rome Air Development Center optical test range in Verona, New York [21]. The use of laser beacons to provide the reference wavefront needed to determine atmospheric turbulence was first demonstrated in 1989 at Kirtland AFB's Starfire Optical Range [21]. The result of many years of development is that AO has been advanced to the point of being a relatively mature field, able to give good performance under conditions of weak atmospheric turbulence.

2.2 Wavefront Sensors

The first step in correcting for the effects of atmospheric turbulence is to measure these effects. This requires a sensor which can measure an optical wavefront.

Light, when treated as a wave, is represented in an x, y plane as the optical field $U(x, y, t) = A(x, y, t)\exp[j\phi(x, y, t) \pm \omega t]$ where $A(x, y, t)$ represents the amplitude of the wave field, $\phi(x, y, t)$ is the phase and ω is the angular frequency of the field [22]. Considering only coherent light, we drop the ωt dependence and simply consider the phasor form of the field $U(x, y) = A(x, y)\exp[j\phi(x, y)]$.

Measuring an optical field is difficult because the period of optical waves is on the order of femtoseconds, much too fast to detect directly. As such, there is no such thing as

a field detector. Instead, sensors able to detect the intensity of light $I(x, y, t) = \|U(x, y, t)\|$ are utilized. The field of the light must be inferred from these intensity measurements.

The field intensity can be measured by media such as the retina of an eye, photographic film or solid-state photo-detector arrays. All these media work effectively the same in that they integrate the intensity of the field over some finite amount of time. This can be represented by

$$I(x, y, t_k) = C \int_{t_{k-1}}^{t_k} \|U(x, y, t)\|^2 dt = C \int_{t_{k-1}}^{t_k} A^2(x, y, t) dt \quad (1)$$

where t_k is the nomenclature for the k^{th} time interval from t_{k-1} to t_k and C converts the units of $\|U(x, y, t)\|$ and $A^2(x, y, t)$ (watts/m²) to the units of intensity (Joules/m²) [20].

A photo-detector array is particularly relevant for AO as it is the medium used to measure the field in AO systems. Since the pixel size of a photo-detector array is fixed and finite, the formula becomes

$$I(x_i, y_j, t_k) = \int_{y_j} \int_{x_i} I(x, y, t_k) dx dy \quad (2)$$

where x and y are integrated over the area of the ij^{th} detector pixel. $I(x_i, y_j, t_k)$ is then the average intensity of the ij^{th} pixel for the k^{th} time frame from t_{k-1} to t_k . The key point here is that $I(x_i, y_j, t_k)$ is a discrete representation of $I(x, y, t)$. If pixel sizes are kept small and time frames kept short, the representation is generally adequate.

Having the intensity of the field, the discrete amplitude $A(x_i, y_j, t_k)$ is easy to determine as $\sqrt{I(x_i, y_j, t_k)}$ [19]. Determining the phase of the field $\phi(x_i, y_j, t_k)$ is more problematic.

2.2.1 Shack-Hartmann Wavefront Sensor. The most common wavefront gradient sensor is the Shack-Hartmann WFS. As shown in Figure 7, this sensor works by using a lenslet array to divide the aperture into multiple sub-apertures [11]. Each sub-aperture's lenslet focuses incoming light onto a focal-plane detector. The detector has a small grid of photo-detector pixels assigned to each sub-aperture. If the light in one sub-aperture is

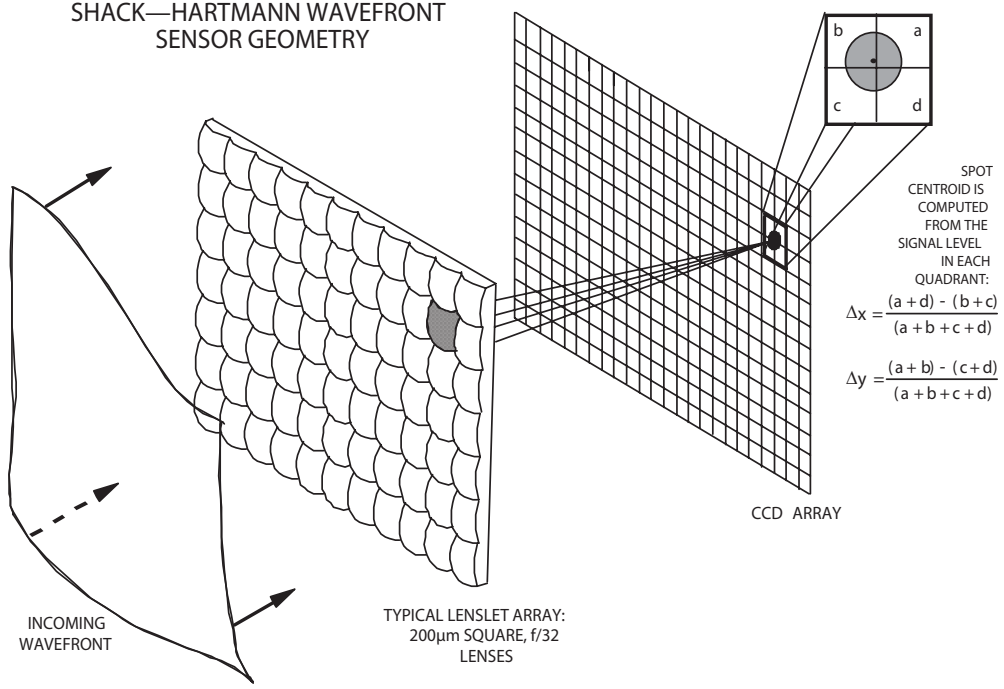


Figure 7: Shack-Hartmann lenslet diagram [40]

propagating in a planar fashion, the light is focused in the center of that sub-aperture's photo-detector grid. If the light entering a sub-aperture has a non-zero average tilt (or average phase gradient), then the light is focused off-center on the receiving array. The x and y centroid of the received light is computed by performing a weighted average of the intensity of the light received by each pixel assigned to a sub-array in the equation

$$\bar{x}_{ij} = \frac{\sum_{m=1}^N \sum_{n=1}^N \bar{x}_{mn} I(x_m, y_n, t_k)}{\sum_{m=1}^N \sum_{n=1}^N I(x_m, y_n, t_k)} \quad (3)$$

and

$$\bar{y}_{ij} = \frac{\sum_{m=1}^N \sum_{n=1}^N \bar{y}_{mn} I(x_m, y_n, t_k)}{\sum_{m=1}^N \sum_{n=1}^N I(x_m, y_n, t_k)} \quad (4)$$

where N is the number of pixels in the x and y directions (assumed to be the same) of the ij^{th} sub-aperture and $\bar{x}_{mn}$ and $\bar{y}_{mn}$ are the x and y center of the mn^{th} pixel in the sub-

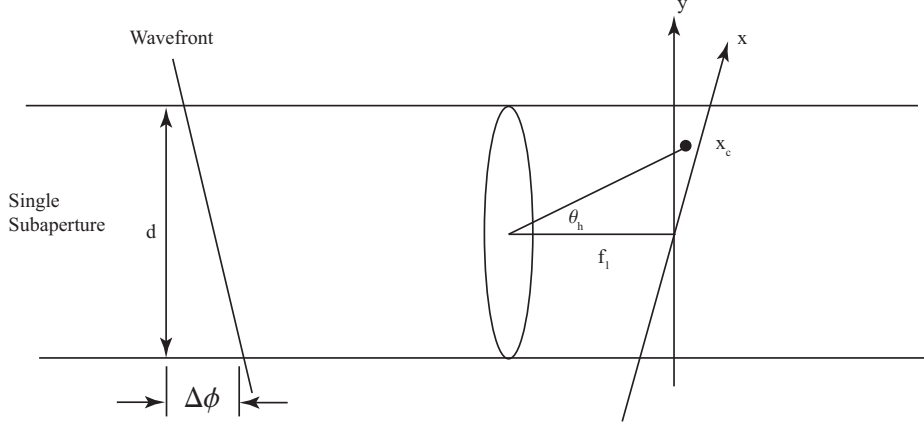


Figure 8: Determining phase tilt from a S-H WFS [8]

aperture. Note that for Figure 7, $N = 2$ with detector pixels symmetric about the optical axis of the subaperture. This simplifies Equations (3) and (4) to the expressions included in the figure where a , b , c and d are the intensities $I(x_m, y_n, t_k)$ and the resultant Δx and Δy are the centroid of the ij^{th} sub-aperture $\bar{x}_{ij}$ and $\bar{y}_{ij}$.

As shown in Figure 8, the average tilt of the field over the sub-aperture in the x and y directions can be determined by the geometry of the setup. The phase gradients in the x and y directions can be calculated as

$$\left(\frac{\partial\phi}{\partial x}\right)_{ij} = \frac{\bar{x}_{ij}}{2\pi\lambda f} \quad (5)$$

and

$$\left(\frac{\partial\phi}{\partial y}\right)_{ij} = \frac{\bar{y}_{ij}}{2\pi\lambda f} \quad (6)$$

where f is the focal length of the sub-aperture lenslet and λ is the wavelength of the light.

The net result of these gradient calculations is that arrays of phase gradients, $(\frac{\partial\phi}{\partial x})_{ij}$ and $(\frac{\partial\phi}{\partial y})_{ij}$, are generated and can be passed on as wavefront information to a controller. The controller can reconstruct the field from those gradient measurements.

2.2.2 Temporally Phase-Shifted Self-Referencing Interferometer. As shown in Figure 9, a self-referencing interferometer (SRI) takes a different approach to measuring wavefront information. Incoming light is split into two separate beams. The reference beam is created by focusing one leg into a single mode optical fiber. This strips all but the DC,

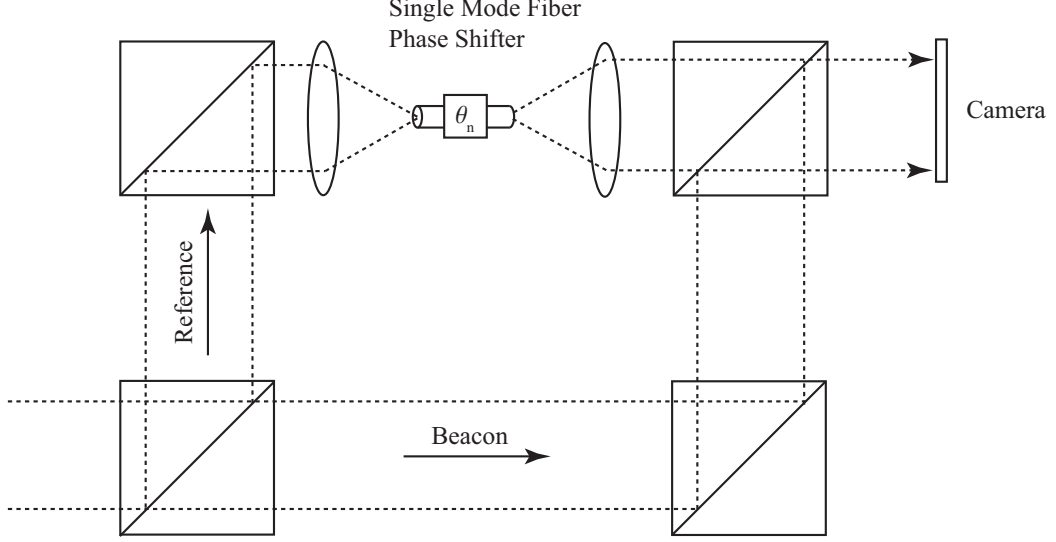


Figure 9: Conceptual diagram of a temporally phase-shifted SRI [35]

or planar, component of the light by spatially filtering it into the fiber's single mode. The light that exits the fiber acts as a point source. After the fiber, the light is collimated by a second lens. At this point, it is a uniform plane wave while retaining coherence to the beacon (which is required to create an interferogram). Note that the fiber acts not only as a spatial filter, but as a phase shifter with an induced phase shift θ_n . The relative phase shift θ_n of the fiber can be varied with an optical phase shifter which slightly stretches the fiber in order to alter the phase of the outgoing light. The ability to control and vary the relative phase of the reference beam is critical to determining the field of the beacon signal.

The reference beam optical field can be expressed as

$$U_r(t) = A_r(t)e^{-j\theta_n} \quad (7)$$

where $A_r(t)$ is the amplitude of the reference beam and θ_n is a variable that can be controlled by the SRI's phase shifter. Note the lack of spatial dependence in $U_r(t)$.

The SRI recombines the beacon and reference beams expressed as

$$U_{total}(x, y, t) = U_b(x, y, t) + U_r(t) = A(x, y, t)e^{j\phi(x, y, t)} + A_r(t)e^{j\theta_n} \quad (8)$$

This combination creates interference patterns between the two beams. Wherever the beacon beam is in phase with the reference beam, there is constructive interference and wherever the beacon is out of phase with the reference there is destructive interference [26].

The intensity of the interference pattern at the photo-detector array is

$$\begin{aligned}
I_n(x, y, t) &= \|U_{total}(x, y, t)\|^2 = U_{total}U_{total}^* \\
&= (U_b + U_r)(U_b + U_r)^* \\
&= U_bU_b^* + U_rU_r^* + U_bU_r^* + U_rU_b^* \\
&= \|U_b\|^2 + \|U_r\|^2 + A(x, y, t)A_r(e^{j[\phi(x, y, t) - \theta_n]} + e^{-j[\phi(x, y, t) - \theta_n]}) \\
&= I_b(x, y, t) + I_r(x, y, t) + 2\sqrt{I_b(x, y, t)I_r(x, y, t)}\cos[\phi(x, y, t) - \theta_n]
\end{aligned} \tag{9}$$

The photo detector array measures the interference pattern, recording the data for each pixel $I_n(x_i, y_j, t_k)$ as in Equation (2) and passes the array of discretized data $I_n(t_k)$ to the controller for processing.

In the interference pattern $I_n(t_k)$, the subscript n indicates explicit correspondence to the reference phase delay θ_n . Chosen properly, each different phase offset will create an independent interference pattern when the reference and beacon beams are combined. At least three independent interference patterns are required to determine the three unknowns of amplitude $A(x_i, y_j, t_k)$, phase $\phi(x_i, y_j, t_k)$ and reference amplitude $A_r(x_i, y_j, t_k)$. Four interference patterns are more typically used in a pattern where $\theta_1 = 0$, $\theta_2 = \pi/2$, $\theta_3 = \pi$ and $\theta_4 = 3\pi/2$ for $n=1, 2, 3$ and 4 so that a four frame sequence of interference patterns would be

$$\begin{aligned}
I_1(x_i, y_j, t_{k-3}) &= I_{\text{bcn}}(x_i, y_j, t_{k-3}) + I_r(x_i, y_j, t_{k-3}) \\
&\quad + 2\sqrt{I_{\text{bcn}}(x_i, y_j, t_{k-3})I_r(x_i, y_j, t_{k-3})} \cos(\phi(x_i, y_j, t_{k-3})) \\
I_2(x_i, y_j, t_{k-2}) &= I_{\text{bcn}}(x_i, y_j, t_{k-2}) + I_r(x_i, y_j, t_{k-2}) \\
&\quad + 2\sqrt{I_{\text{bcn}}(x_i, y_j, t_{k-2})I_r(x_i, y_j, t_{k-2})} \cos(\phi(x_i, y_j, t_{k-2}) - \pi/2) \\
I_3(x_i, y_j, t_{k-1}) &= I_{\text{bcn}}(x_i, y_j, t_{k-1}) + I_r(x_i, y_j, t_{k-1}) \\
&\quad + 2\sqrt{I_{\text{bcn}}(x_i, y_j, t_{k-1})I_r(x_i, y_j, t_{k+2})} \cos(\phi(x_i, y_j, t_{k-1}) - \pi) \\
I_4(x_i, y_j, t_k) &= I_{\text{bcn}}(x_i, y_j, t_k) + I_r(x_i, y_j, t_k) \\
&\quad + 2\sqrt{I_{\text{bcn}}(x_i, y_j, t_k)I_r(x_i, y_j, t_k)} \cos(\phi(x_i, y_j, t_k) - 3\pi/2)
\end{aligned} \tag{10}$$

where $I_{\text{bcn}}()$ is the intensity of the beacon and $I_r()$ is the intensity of the reference beam.

Then, making the assumption that field is being sampled fast enough that the field has not changed between the four successive interference patterns,

$$\begin{aligned}
I_1(x_i, y_j, t_k) &= I_1(x_i, y_j, t_{k-3}) \\
I_2(x_i, y_j, t_k) &= I_2(x_i, y_j, t_{k-2}) \\
I_3(x_i, y_j, t_k) &= I_3(x_i, y_j, t_{k-1})
\end{aligned}$$

so that the field can be determined [24] as

$$\begin{aligned}
U(x_i, y_j, t_k) &= \frac{1}{4\sqrt{I_r}}(I_1(x_i, y_j, t_k) - I_3(x_i, y_j, t_k) + j[I_2(x_i, y_j, t_k) - I_4(x_i, y_j, t_k)]) \\
&= \frac{1}{2}\sqrt{I_{\text{bcn}}(x_i, y_j, t_k)}[\cos(\phi(x_i, y_k, t_k)) + j\sin(\phi(x_i, y_j, t_k))] \\
&= A_{\text{bcn}}(x_i, y_j, t_k)e^{j\phi(x_i, y_j, t_k)}
\end{aligned} \tag{11}$$

From Equation (11), the average phase of the field in the ij^{th} pixel can be determined as

$$\phi(x_i, y_j, t_k) = \left\{ \begin{array}{ll} \tan^{-1} \left[\frac{I_2(x_i, y_j, t_k) - I_4(x_i, y_j, t_k)}{I_1(x_i, y_j, t_k) - I_3(x_i, y_j, t_k)} \right] & \forall I_1(x_i, y_j, t_k) \neq I_3(x_i, y_j, t_k) \\ \frac{\pi}{2} & \forall I_2(x_i, y_j, t_k) > I_4(x_i, y_j, t_k), \\ & I_1(x_i, y_j, t_k) = I_3(x_i, y_j, t_k) \\ -\frac{\pi}{2} & \forall I_2(x_i, y_j, t_k) < I_4(x_i, y_j, t_k), \\ & I_1(x_i, y_j, t_k) = I_3(x_i, y_j, t_k) \\ \text{Undefined} & \forall I_2(x_i, y_j, t_k) = I_4(x_i, y_j, t_k), \\ & I_1(x_i, y_j, t_k) = I_3(x_i, y_j, t_k) \end{array} \right. \quad (12)$$

Being able to derive the field's average phase for each of the pixels of the photo-detector array makes each pixel equivalent to a subaperture in a gradient WFS which uses at least four, and often sixteen detector pixels. Unlike the case of a gradient sensor, however, an SRI requires no further reconstruction because the field is already computed through simple trigonometry.

The phase of the reference beam is controlled so that it shifts between successive frames. This is depicted in Figure 11. The advantage of the temporally phase-shifted SRI is the relative simplicity of its design [10]. The disadvantage is that the system assumes that the turbulence is unchanged while the temporal interference pattern is generated. As a result, the frame rate must be kept very high. The Greenwood frequency f_G , which can be determined from turbulence and wind profiles, describes how fast the field is changing due to the dynamic nature of the turbulence [1]. In order to be valid, the frame rate of a temporally designed SRI must be much higher than the Greenwood frequency of the turbulence being encountered for the field estimations to be accurate [38].

An extended Kalman filter (EKF) (see Section 2.4.2.2) has been used to estimate the field from the individual interferograms produced by a temporally phase-shifted SRI [41]. This improves performance at lower sampling rates, allowing effective operation of

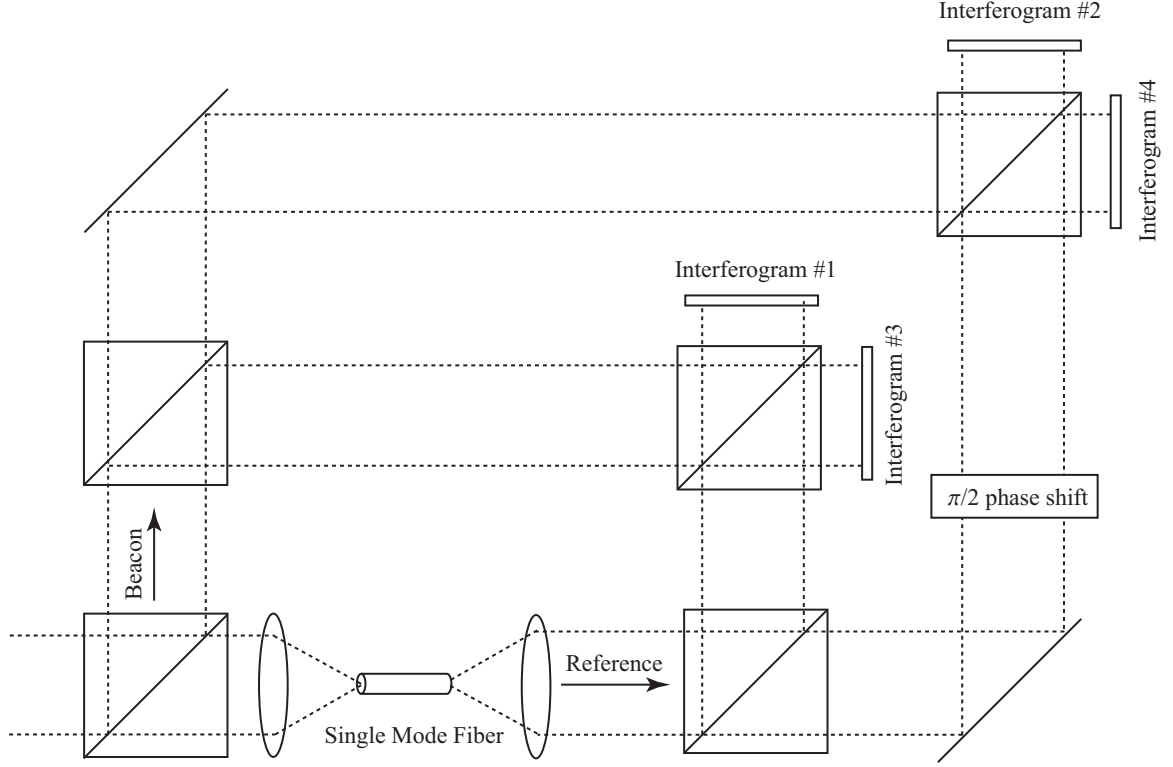


Figure 10: Conceptual Spatially Designed SRI [37]

temporally phase-shifted SRI designs under a broader range of turbulence conditions, in particular under increased Greenwood frequency conditions.

2.2.3 Spatial SRI. A spatially phase-shifted SRI is similar to a temporal SRI except that instead of creating multiple reference beams by shifting the phase of the reference beam temporally, a spatially phase-shifted SRI splits the incoming light into eight parts - four beacon and four reference. Each of the four reference beams have different optical paths and are shifted from each other by 0 , $\pi/2$, π , and $3\pi/2$ as shown in Figure 10. Each of the four different reference beam paths is combined with a beacon path and then all four interference patterns are measured simultaneously.

Figure 12 depicts all four interference patterns on a single photo-detector. The four measurements are then processed into an array of phases via Equation (12). The advantage of a spatially phase-shifted SRI is that all four measurements are taken simultaneously. The disadvantages are the added complexity of the additional optical paths, the difficulty in

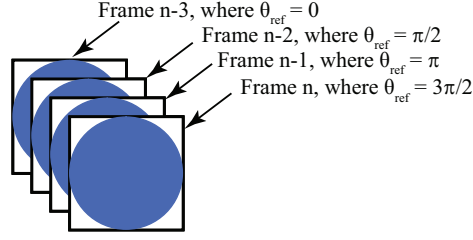


Figure 11: Frame sequence for a temporal SRI [38]. The black outline indicates the extent of the camera. The blue circle is the area of the pupil interferogram.

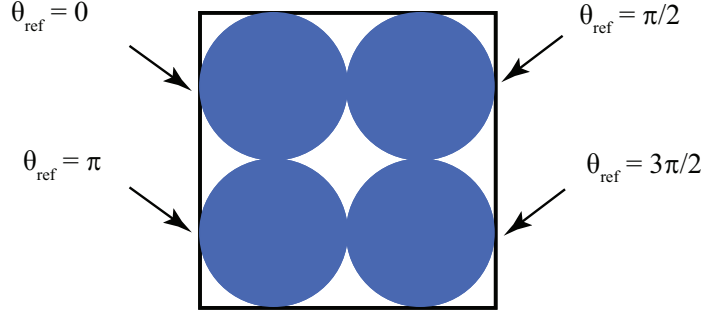


Figure 12: Spatial SRI [34]

getting everything aligned properly, and the loss of intensity caused by dividing the incoming beam into eight parts instead of two [10].

2.2.4 SRI performance. The accuracy of the SRI WFS is characterized by the estimation Strehl ratio S_N

$$S_N = \frac{1}{1 + \text{SNR}_M^2} \quad (13)$$

where SNR_M is the modulation signal-to-noise ratio. SNR_M is the key quantity for SRI WFS performance [5]. For a simplistic model SNR_M is defined as

$$\frac{(\frac{4}{N_{\text{split}}})^2 I_r I_s}{(\frac{4}{N_{\text{split}}})^2 I_r I_s + 4(\sigma_{rd}^2 + \sigma_q^2 + \sigma_{\text{shot}}^2)} \quad (14)$$

where σ_{rd}^2 is the variance of the camera read noise, σ_q^2 is the variance of the quantization noise, and σ_{shot}^2 is the variance of the shot noise [37]. N_{split} is the number of branches that the optical field is split into. For a temporally-shifted SRI $N_{\text{split}} = 2$, while for a spatially-shifted SRI $N_{\text{split}} = 4$.

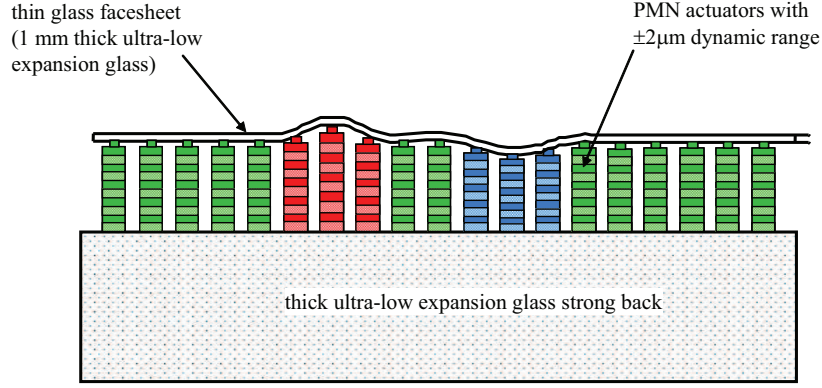


Figure 13: Continuous Facesheet DM [40]

2.3 Wavefront Correction Devices

There are many different wavefront correction devices which all control the wavefront by varying different portions of the optical path. DMs are the only type of corrector discussed here as they are the device intended for use in this project.

2.3.1 Continuous Facesheet Deformable Mirrors. A DM as depicted in Figure 13 is an array of actuators or pistons covered by a sheet of flexible reflective material. The actuators deform the mirror so that parts of the mirror are higher than others, making the optical path shorter for the higher parts. The surface of the mirror is continuous, causing a smooth transition for the phase corrections. Using a continuous facesheet DM has both advantages and disadvantages. An advantage is that the DM nicely smooths out the discrete solution created by a digital controller. A disadvantage is that the DM cannot make the instantaneous transitions necessary to compensate for areas in the field which have phase discontinuities (phase discontinuities are covered in Section 2.6).

2.3.2 Segmented Deformable Mirrors. Segmented mirrors have individual mirror segments controlled by piston-like actuators as in Figure 14. The segments are raised and lowered and create a constant phase correction to that area of the mirror. This is a drawback in that only the first Zernike mode (piston) over the segment is corrected and any tip/tilt or higher-order Zernike modes that a field may have over a segment is uncorrected. There is also some performance loss due to the diffraction effects of the edges of all the mirror segments as well as some small loss due to the spacing between mirror segments. The

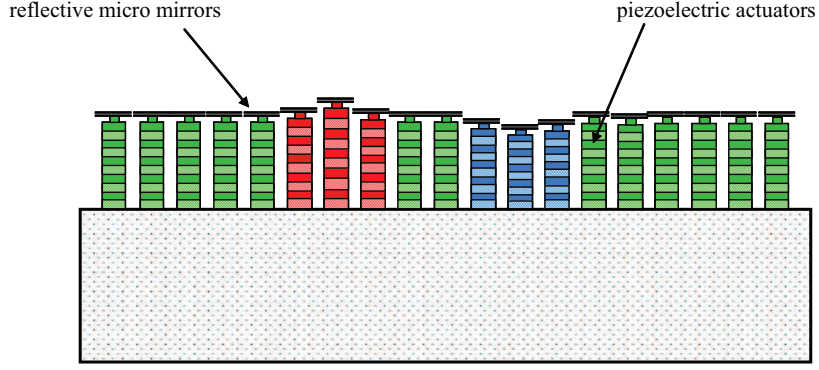


Figure 14: Segmented DM [40]

advantage of a segmented mirror, however, is in its ability to emulate discontinuous phase shifts. The correction applied to the field does not have to be unwrapped since a segmented mirror handles the unnecessary 2π shifts of an unwrapped field without a problem. This makes a segmented mirror very simple to implement with an SRI which gives the phases of a field. Using deterministic control, the phases are simply scaled by a gain and applied to the segmented mirror.

Segmented mirrors do exist that have multiple actuators per segment allowing the segments to tilt. This allows the mirror to correct the tip and tilt of the field over a specific segment, but higher order Zernike modes are not addressed and the adverse diffraction effects of the segment edges are encountered. In this case, the tilt of the field over a segment would have to be determined and appropriate controls applied to the mirror.

2.4 Stochastic Control

Control is the heart of an AO system. It does not matter how good the WFS or wavefront corrector is if the controller connecting the two is insufficient. Conversely, while a good controller cannot make up for an inadequate WFS or wavefront corrector, it can at least determine the best control possible given the limitations of the other elements of the system.

The controllers of conventional AO control systems have been essentially deterministic in that they get input from the WFS, multiply this input by a gain, and then apply that output to the wavefront corrector. Beyond trying to minimize certain noise effects, they

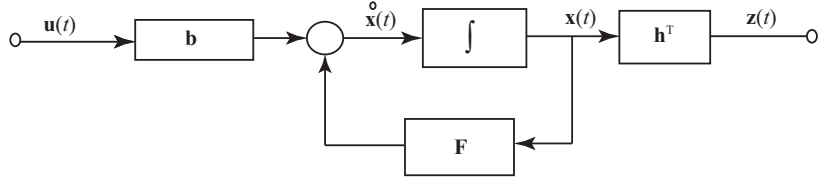


Figure 15: State space system representation [28]

make no attempt to treat the system as stochastic. The major element of this project is to consider the AO system as a stochastic process and control the system accordingly.

“In the mathematics of probability, a stochastic process or random process is a process that can be described by a probability distribution.” [46] In trying to control a stochastic process, it is reasonable to divide problems into modeling, estimation, and control. While adequate coverage of the subject would require rewriting the textbooks, a brief description of the pertinent background material is presented.

2.4.1 State Space System Representation. Systems are modeled using state variables. This is shown in Figure 15, where $\mathbf{x}(t)$ is the state space vector, the value of which is known as the *state* of the system. The array $\mathbf{B}$ describes how the control inputs $\mathbf{u}(t)$ affect the states, $\mathbf{H}^T$ models the relationship between the states and the output $\mathbf{z}(t)$, and $\mathbf{F}$ describes the linear dynamics of the system. States of the system do not have to be known identifiable parameters of the system (although they usually are); they simply must be a set of variables which adequately describe the system behavior [31]. Moreover, there can be several state spaces, all of which adequately describe the system. Designers can choose which state space is most amenable to the design process.

2.4.2 Stochastic Modeling. The key to good control design is developing an accurate model of the system to be controlled. In stochastic modeling, models are built incorporating both the deterministic physical system as well as the stochastic noise influences on that system. In the adaptive optics model, the noise sources are the effect of atmospheric turbulence and the other sources of noise identifiable in the system such as shot and detector noise. These noise sources should be characterized based on existing data and validated physics in order to create the best model possible.

2.4.2.1 Gauss-Markov State Variables. As part of a stochastic system, a state variable will never be known, but instead have a probability density function (PDF) which can be estimated. Designers strive to choose state variables in such a way that they are Gauss-Markov. Being Gaussian allows the PDF of the variable to be fully described by a mean and variance. Being Markov implies that the future behavior of a variable is only dependent on its previous state, relieving the system from maintaining the entire time series of previous state information.

Once chosen, the state variables are vectorized into an N -dimensional vector $\mathbf{x}(t)$ where N is the number of state variables and the bold typeface indicates the vector nature of $\mathbf{x}(t)$. The PDF of $\mathbf{x}(t)$ is described by $\hat{\mathbf{x}}(t)$ and $\mathbf{P}(t)$, where $\hat{\mathbf{x}}(t)$ is the N -dimensional best estimate of $\mathbf{x}(t)$ (or mean, since it is Gaussian) and $\mathbf{P}(t)$ is the $N \times N$ state covariance matrix. In the covariance matrix, the diagonal terms are the variances of each variable, and the non-diagonal terms are the covariances between different variables.

2.4.2.2 Model types. The system is modeled as the continuous-time differential equation $\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t)$ so that the states' dynamics are functions of the states, inputs $\mathbf{u}(t)$, system dynamic noises $\mathbf{w}(t)$ and time t . The output of the system is a function of the state variables and time, $\mathbf{z}(t) = h(\mathbf{x}(t), t)$. All systems are modeled as either linear or non-linear. For the systems considered for this project, we will restrict ourselves to linear time-invariant (LTI) models or non-linear time-invariant models that can be readily linearized. However, investigation into non-linear effects is conducted only as needed.

Linear systems and Kalman filters. If the system is adequately represented with LTI models, then the system is well represented by the differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{G}\mathbf{w}(t) \quad (15)$$

where $\mathbf{F}$ represents how the states change over time, $\mathbf{B}$ represents how the inputs affect the states and $\mathbf{G}$ represents how the system noises affect the states. Given that noisy measurements are available from sensors, the output of the system becomes

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t) \quad (16)$$

where $\mathbf{H}(t)$ determines the output from the states and $\mathbf{v}(t)$ is measurement noise.

The solution to Equation (15) is

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0 + \int_{t_0}^t \Phi(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau + \int_{t_0}^t \Phi(t, \tau)\mathbf{G}(\tau)d\beta(\tau) \quad (17)$$

where $\Phi(t_2, t_1)$ is the state transition matrix which describes how the system states change from time t_1 to time t_2 . Details of this are left to the applicable text [28]

Equivalent discrete-time system model. In order to implement the system models using digital computers, an equivalent discrete-time system model must be developed from the continuous-time model. Again, details are well covered applicable texts [28], but the result is

$$\mathbf{x}(t_{i+1}) = \Phi(t_{i+1}, t_i)\mathbf{x}(t_i) + \left[\int_{t_i}^{t_{i+1}} \Phi(t_{i+1}, \tau)\mathbf{B}(\tau)d\tau \right] \mathbf{u}(t_i) + \int_{t_i}^{t_{i+1}} \Phi(t_{i+1}, \tau)\mathbf{G}(\tau)d\beta(\tau) \quad (18)$$

$$\mathbf{x}(t_{i+1}) = \Phi(t_{i+1}, t_i)\mathbf{x}(t_i) + \mathbf{B}_d(t_i)\mathbf{u}(t_i) + \mathbf{w}_d(t_i) \quad (19)$$

where $\mathbf{B}_d(t_i)$ is defined as $\mathbf{B}_d(t_i) = \int_{t_i}^{t_{i+1}} \Phi(t_{i+1}, \tau)\mathbf{B}(\tau)d\tau$ and $w_d(t_i)$ is a zero-mean discrete white-Gaussian process with variance Q_d .

Linear Kalman Filter. From the equivalent discrete-time model, a recursive estimation method known as a Kalman filter can be developed for discrete-time systems. The development and solution to the optimal linear Kalman filter is well documented in reference [28] and is only summarized here.

A discrete-time Kalman filter is illustrated in Figure 16 and has two parts. The first is the propagation of the state variable estimate vector and covariance matrix between measurements. The second is the updating of the state variable estimate and covariance matrix when measurements are taken.

In the propagation portion of the filter, the state estimate vector is propagated one frame into the future by the state transition matrix $\Phi(t_i, t_{i-1})$ and combined with the effects of any system inputs $\mathbf{u}(t)$. The state transition matrix and the dynamic noise characteristics

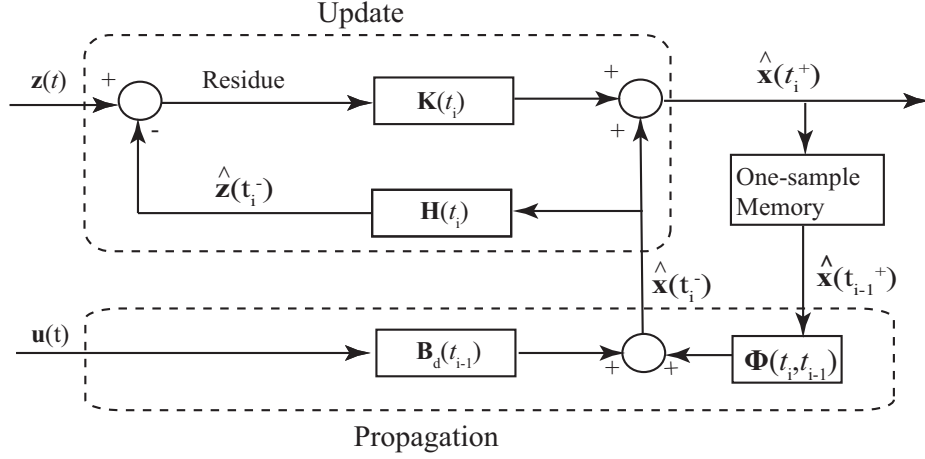


Figure 16: Kalman filter illustration [30]

control the propagation of the covariance matrix. The equations for this are

$$\begin{aligned}\hat{\mathbf{x}}(t_i^-) &= \Phi(t_i, t_{i-1})\hat{\mathbf{x}}(t_{i-1}^+) + \mathbf{B}_d(t_{i-1})\mathbf{u}(t_{i-1}) \\ \mathbf{P}(t_i^-) &= \Phi(t_i, t_{i-1})\mathbf{P}(t_{i-1}^+)\Phi^T(t_i, t_{i-1}) + \mathbf{G}_d(t_{i-1})\mathbf{Q}_d(t_{i-1})\mathbf{G}_d^T(t_{i-1})\end{aligned}$$

where the ‘-’ and ‘+’ superscripts indicate times just before and after an update within the Kalman filter. Here $\mathbf{G}_d(t)$ is the matrix representing how the system noises affect the covariance matrix $\mathbf{P}$.

In the update portion of the filter, a residue is formed by subtracting the best estimate of what the measurement $\hat{\mathbf{z}}(t)$ from the measurement itself $\mathbf{z}(t)$. The updated measurement is then formed by summing the old estimate $\hat{\mathbf{x}}(t_i^-)$ and the residue weighted by the Kalman gain $\mathbf{K}(t_i)$, a pre-computed matrix based on the covariance matrices of the old estimate and new measurements. The Kalman gain matrix also controls the updating of the state covariance matrix. The equations for this are

$$\mathbf{K}(t_i) = \mathbf{P}(t_i^-)\mathbf{H}^T(t_i)[\mathbf{H}(t_i)\mathbf{P}(t_i^-)\mathbf{H}^T(t_i) + \mathbf{R}(t_i)]^{-1} \quad (20)$$

$$\hat{\mathbf{x}}(t_i^+) = \hat{\mathbf{x}}(t_i^-) + \mathbf{K}(t_i)[\mathbf{z}(t_i) - \mathbf{H}(t_i)\hat{\mathbf{x}}(t_i^-)] \quad (21)$$

$$\mathbf{P}(t_i^+) = \mathbf{P}(t_i^-) - \mathbf{K}(t_i)\mathbf{H}(t_i)\mathbf{P}(t_i^-) \quad (22)$$

where $\mathbf{H}(t)$ is the matrix that converts the best estimate of the state of the system to the best estimate of the output being measured $\hat{\mathbf{z}}(t) = \mathbf{H}(t)\hat{\mathbf{x}}(t)$, $\mathbf{R}(t)$ is the covariance matrix of measurement noises and $\mathbf{z}(t_i)$ is the measurement for the i^{th} frame.

The implementation of a Kalman filter often simplifies in practice. In uncontrolled systems, $\mathbf{u}(t) = 0$. A fixed frame rate and time-invariant system dynamics make $\Phi(t_i, t_{i-1})$, and $Q_d(t_i)$ temporally constant. The measurement matrix $\mathbf{H}(t)$ is usually time-invariant. Sometimes $\mathbf{R}(t)$ is time-invariant, which tremendously simplifies matters since, in steady-state, $\mathbf{K}(t_i)$ is constant and pre-computable. In the simplest case where all these things are true, a Kalman filter reduces to simply taking a weighted average of the current estimate and information from new measurements at each update.

Non-Linear systems. Given that a goal of this research is to determine a valid model for the system, it may be determined that nonlinear models are required. In this case, either nonlinear filtering, or more likely, linearized filtering such as the extended Kalman filter (EKF) would be appropriate [29].

An EKF works largely the same as a linear Kalman filter, except $\mathbf{F}$ and $\mathbf{H}$ have to be recalculated as

$$\begin{aligned} \mathbf{H}[t_i; \hat{\mathbf{x}}(t_i^-)] &\equiv \left. \frac{\partial \mathbf{h}[\mathbf{x}(t_i), t_i]}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(t_i^-)} \\ \text{and } \mathbf{F}[t; \hat{\mathbf{x}}(t/t_i)] &\equiv \left. \frac{\partial \mathbf{f}[\mathbf{x}, \mathbf{u}(t), t]}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(t/t_i)} \end{aligned}$$

to account for the non-linearities. Development, solution and application of EKFs are covered in [29].

2.4.3 Stochastic Estimation. Stochastic estimation is the process of estimating the state variables of the system based on an adequate system model. Kalman filters are the most well-known stochastic estimators which optimally combine new measurements with current estimates of state variables to determine an updated estimate of these state variables.

2.4.4 Stochastic Control. In contrast to deterministic control methods, stochastic control methods provide robust solutions in the presence of system uncertainty. Improvements in stability and performance can be achieved by adequately accounting for the unmodeled effects on either the system dynamics or measurement devices.

2.4.5 LQG. The design process of developing a stochastic controller may be intractable without some bounds. This involves properly defining what makes a system good or bad in terms of a system “cost,” under what conditions this “cost” should be minimized, what kinds of system models are used and what types of noise sources are assumed. Often “cost” is intuitive, but it needs to be defined in order to regiment the design process and determine what “optimality” means for the problem at hand.

In a stochastic control design, representing the system as *linear*, costs as *quadratic*, and noises (both system and measurement) as *Gaussian* (LQG) has proven to be of great benefit. An LQG system requires that the system model be linear, have a quadratic cost structure for the control and have Gaussian noise inputs [30].

A quadratic cost function takes the form

$$J = E \left[\sum_{i=0}^N [\mathbf{x}^T(t_i) \mathbf{X}(t_i) \mathbf{x}(t_i) + \mathbf{u}^T(t_i) \mathbf{\Gamma}(t_i) \mathbf{u}(t_i)] + \mathbf{x}^T(t_{N+1}) \mathbf{X}_f \mathbf{x}(t_{N+1}) \right]$$

where $\mathbf{X}(t_i)$ is a weighting matrix defining the relative cost associated with various states and $\mathbf{\Gamma}(t_i)$ is a weighting matrix defining the cost of various control inputs [30]. The quadratic nature of the cost function is in the squared terms of both states and control inputs.

The primary benefit of satisfying the LQG assumptions is a tremendous simplification of the mathematics of the optimization process. The simplified mathematics allow a regimented and disciplined design process. While convenience for the designer is appealing, the assumptions must also be adequate, and the design process must yield quality products. Thankfully, most systems can be adequately represented by a linear model, most design requirements can be adequately reflected as quadratic costs and most noise sources are reasonably Gaussian (or can be represented by linear shaping filters and thus satisfying these

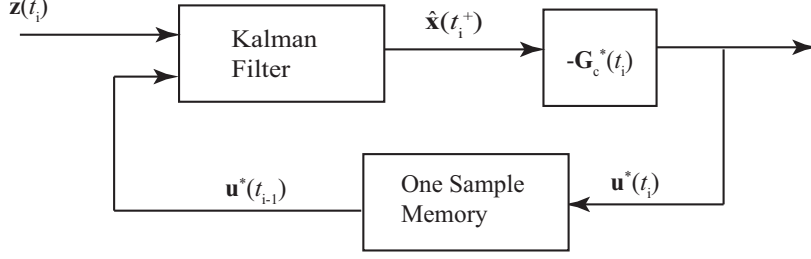


Figure 17: LQG controller illustration [30]

same assumptions). Even more importantly, LQG designs usually work well and produce exceptional results.

Given that the LQG assumptions are valid and stochastic estimation has been used to provide the controller with optimal state estimates, the *certainty equivalence* property can be applied resulting in a *deterministic* optimal control solution. “The optimal stochastic controller for a problem described by linear system models driven by white Gaussian noise, subject to a quadratic cost criterion consists of an optimal linear Kalman filter cascaded with the optimal feedback gain matrix of the corresponding *deterministic* optimal control problem.” [30] This simplifies the design process into two steps as shown in Figure 17:

1. Design an optimal stochastic estimator using the Kalman filter
2. Separately design an optimal *deterministic* controller for the original system.

Under LQG assumptions, the optimal feedback for the system is given by

$$\mathbf{u}^*(t_i)[\hat{\mathbf{x}}(t_i^+), t_i] = -\mathbf{G}_c^*(t_i)\hat{\mathbf{x}}(t_i^+), \quad (23)$$

where $\mathbf{G}_c^*$ is the optimal feedback controller gain given by

$$\mathbf{G}_c^* = [\mathbf{\Gamma}(t_i) + \mathbf{B}_d^T(t_i)\mathbf{K}_c(t_{i+1})\mathbf{B}_d(t_i)]^{-1}[\mathbf{B}_d^T(t_i)\mathbf{K}_c(t_{i+1})\mathbf{\Phi}(t_{i+1}, t_i)]. \quad (24)$$

In Equation (24), $\mathbf{\Gamma}$ is a weighting matrix in the cost function of the applied control, $\mathbf{B}_d^T$ is the discretized control matrix and $\mathbf{K}_c$ is the solution to the backward Riccati difference

equation

$$\begin{aligned}\mathbf{K}_c(t_i) &= \mathbf{X}(t_i) + \Phi^T(t_{i+1}, t_i) \mathbf{K}_c(t_{i+1}) \Phi(t_{i+1}, t_i) \\ &\quad - [\Phi(t_{i+1}, t_i) \mathbf{K}_c(t_{i+1}) \mathbf{B}_d(t_i)] [\Gamma(\mathbf{t}_i) + \mathbf{B}_d^T(t_i) \mathbf{K}_c(t_{i+1}) \mathbf{B}_d(t_i)]^{-1} \\ &\quad \times [\mathbf{B}_d(t_i) \mathbf{K}_c(t_{i+1}) \Phi(t_{i+1}, t_i)]\end{aligned}$$

where $\mathbf{X}(t_i)$ is a weighting matrix on state error in the cost function of the system.

In the event that the LQG assumptions are invalid (such as a nonlinear system model), then nonlinear control methods can be employed but at a significant increase in complexity. Alternatively, *assumed* certainty equivalence can be applied that again reduces complexity at the expense of *suboptimal control*.

2.5 Continuity of the Optical Field

Given optical propagation in an unbounded continuous medium with smoothly varying stochastic refractive index, the optical field transverse to the direction of propagation $U(\mathbf{r})$ is described by the *stochastic Helmholtz equation* [1].

$$\nabla^2 U(\mathbf{r}) + k^2 n(\mathbf{r}) U(\mathbf{r}) = 0 \quad (25)$$

where ∇^2 indicates the Laplacian operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$, $U(\mathbf{r})$ is the field, $k = \frac{2\pi}{\lambda}$ where λ is the wavelength of light, $n(\mathbf{r})$ is the index of refraction and $\mathbf{r} = (x, y, z)$ denotes a point in space.

Taking $U(\mathbf{r})$ and $n(\mathbf{r})$ to be defined everywhere, the Laplacian is defined everywhere since Equation (25) can be rearranged to

$$\nabla^2 U(\mathbf{r}) = -k^2 n(\mathbf{r}) U(\mathbf{r}). \quad (26)$$

Since the Laplacian exists, the partial derivatives $\frac{\partial}{\partial x} U(\mathbf{r})$, $\frac{\partial}{\partial y} U(\mathbf{r})$ and $\frac{\partial}{\partial z} U(\mathbf{r})$ must also exist and be continuous.

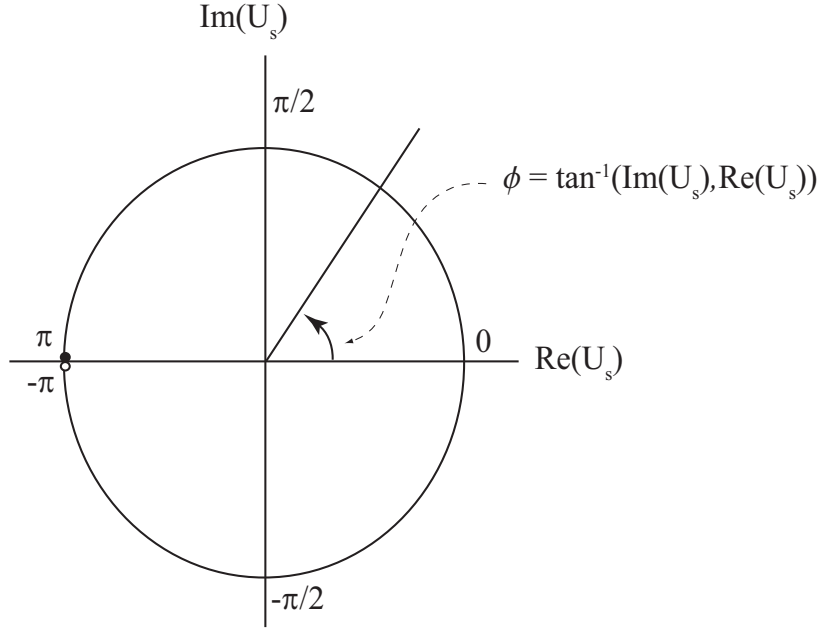


Figure 18: Four quadrant $\tan^{-1}(x, y)$ function. The closed circle at π indicates that $\tan^{-1}(0, x) = \pi \forall x < 0$ while the open circle just below indicates that $\lim_{y \rightarrow 0^-} \tan^{-1}(y, x) = -\pi \forall x < 0$. Thus, the function is discontinuous at $y = 0, x \leq 0$, since $\lim_{y \rightarrow 0^-} \tan^{-1}(y, x) \neq \tan^{-1}(0, x) \forall x < 0$ and $\tan^{-1}(0, 0)$ is undefined.

Since the partial derivatives exist and are continuous, the function $U(\mathbf{r})$ must be continuous. Since $U(\mathbf{r})$ is continuous, the real component $\Re[U(\mathbf{r})]$ and imaginary component $\Im[U(\mathbf{r})]$ of $U(\mathbf{r})$ are continuous.

2.6 Phase Discontinuities

Phase discontinuities arise in two forms: unwrapping and branch points (as discussed in Sections 2.7 and 2.8). These discontinuities arise because of the way phase is defined. Similar to the development of Equation (11), phase is defined as

$$\phi(\mathbf{r}) = \begin{cases} \tan^{-1} \left(\frac{\Im[U(\mathbf{r})]}{\Re[U(\mathbf{r})]} \right) & \forall \Re[U(\mathbf{r})] \neq 0 \\ \frac{\pi}{2} & \forall \Im[U(\mathbf{r})] > 0, \Re[U(\mathbf{r})] = 0 \\ -\frac{\pi}{2} & \forall \Im[U(\mathbf{r})] < 0, \Re[U(\mathbf{r})] = 0 \\ \text{Undefined} & \forall \Im[U(\mathbf{r})] = 0, \Re[U(\mathbf{r})] = 0 \end{cases} \quad (27)$$

where the inverse tangent function is the four-quadrant inverse tangent yielding phases on the range $(-\pi, \pi]$, as shown in Figure 18.

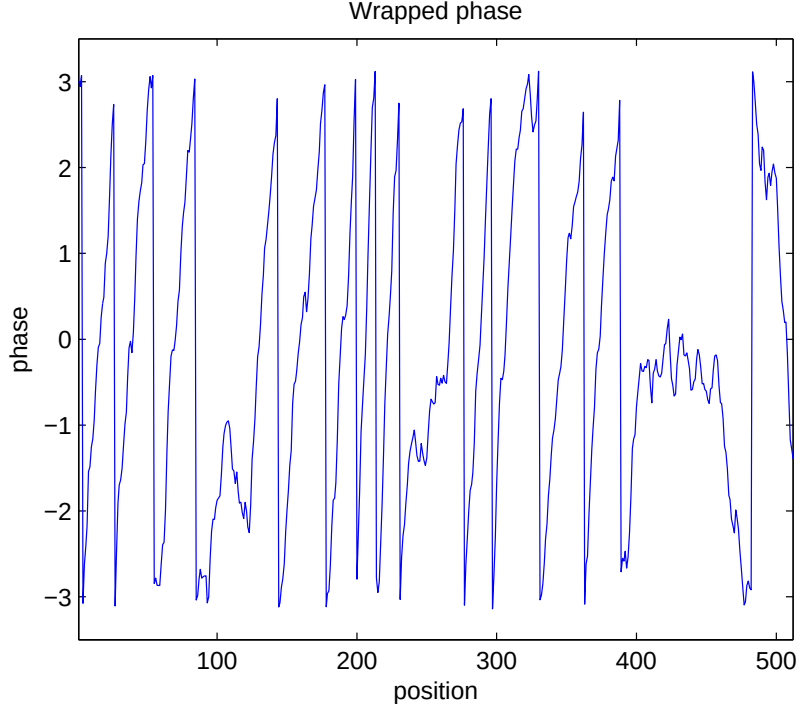


Figure 19: Phases for a single row of a wrapped optical field. The phases are restricted to $(-\pi, \pi]$ from the MATLAB angle function which can be expressed as $\text{angle}(z) = \text{atan2}(\text{imag}(z), \text{real}(z))$ where z is the complex variable given to the function.

2.7 Unwrapping Phase Discontinuities

Unwrapping problems occur where the real portion of the field $\Re[U(\mathbf{r})]$ is negative and the imaginary portion of the field $\Im[U(\mathbf{r})]$ transitions between positive and negative. At these points, the phase jumps 2π . This is a result of the fact that the tangent function is modulo 2π . That is, $\tan(\phi) = \tan(\phi + 2\pi k)$ where k is any integer. This problem can be resolved by adding or subtracting 2π from areas of the field, thus smoothing the field. Figure 19 shows a row of phase angles from a wrapped field, while Figure 20 shows the same row after unwrapping. Figure 21 shows a simplistic unwrapping process (no branch points) in four steps.

2.8 Branch Points

Branch points are an optical phenomenon which is manifested as the curl of the phase $\phi(x, y, t)$ of a light field is non-zero about a point [18]. In mathematical terms, $(\nabla \times (\nabla \phi))(x, y, t) \neq 0$ and $\oint \nabla \phi(x, y, t) \neq 0$ where the closed integral follows a path around

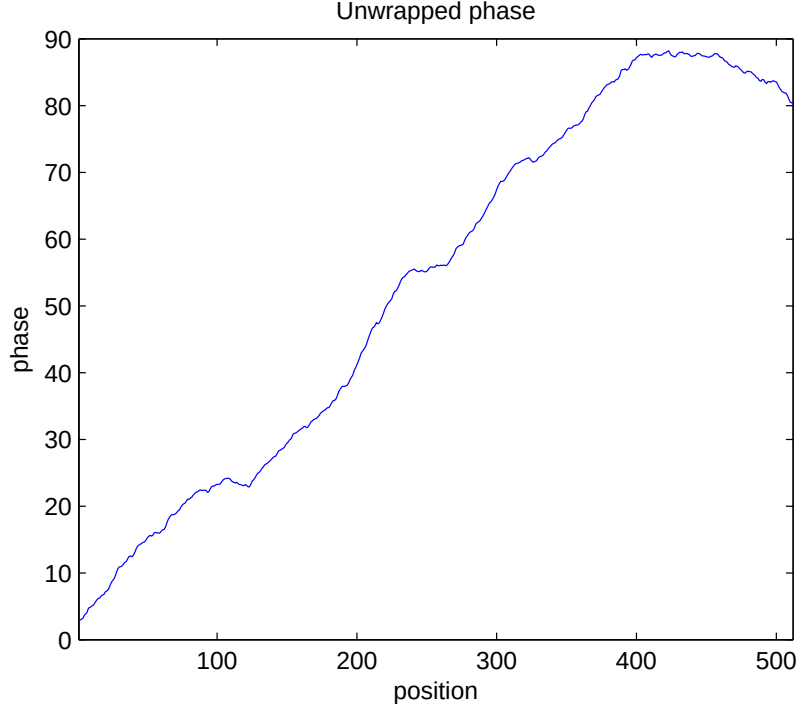


Figure 20: Phases for a single row of an unwrapped optical field. The phases are no longer restricted to $(-\pi, \pi]$. Integer multiples of 2π are added to the phase in order to smooth the phase in a process known as unwrapping.

the single branch point. Thus, starting at a point A and integrating phase differential along a closed path containing a single branch point back to A yields a different phase than at the start. Since point A has not changed, the integration along the closed path will necessarily be some positive or negative integer multiple of 2π . In practice, integrating phase gradients along a path around a point is the way branch points are discovered [14]. When integrating clockwise around a closed path, a positive result indicates a positive branch point. Negative results indicate negative branch points. As such, if a closed loop were to contain both a positive and negative branch point, the effects of the two branch points would have opposite effects and the path integral around the loop would evaluate to zero. Integrations resulting in higher multiples of 2π indicate the presence of multiple branch points of the same polarity. In attempting to detect branch points, the closed paths are kept as small as possible in order to pinpoint their location. Due to the discretization of digital photo-detectors, paths can never be smaller than the size of a single pixel.

2.9 Rotational and Irrotational Field Components

Fields containing branch points are called rotational because there is at least one point with a non-zero curl. The term rotational is appropriate because the gradient of the field forms a vortex, or rotation about the branch point as shown in Figure 22. An irrotational field is one without branch points so that there are not any points of non-zero curl. Such a field would have gradients which do not rotate about any points, as shown in Figure 23. The phase of a rotational field can be divided into rotational and irrotational components [18]. The irrotational component is, of course, irrotational while the rotational component contains all the rotational effects of branch points.

2.10 Difficulties of Branch Points

2.10.1 Wavefront Reconstruction. The presence of a rotational component in the optical field's phase makes effective AO very difficult. First of all, the rotational component is difficult for most systems to detect. Standard systems using a S-H WFS and an LS reconstructor miss the rotational portion of the field. This is the reason the rotational component of the field is sometimes called “the hidden field” [16].

2.10.2 DMs and the irrotational phase. DMs correct the phase of an optical field by altering the surface of a mirror increasing the optical path lengths for some portions of the field and decreasing it for others. The height of the mirror at any point is fixed (for a given time) and the integral of mirror height differentials over a closed path must then equal zero. As a result, the DM is an irrotational surface and well-suited to compensate for irrotational phase errors. To compensate for rotational fields, DMs must use 2π discontinuities where the DM height changes 2π from one pixel to the next in order to implement phase corrections which account for the effects of branch points in the optical field. These discontinuities begin and end at branch points and are known as branch cuts.

2.11 Branch Cuts

To compensate for the non-zero curl of phase differential around branch *points*, a concept of branch *cuts* has been developed. A branch cut is simply an artificially determined line in the detector plane where a discontinuity is forced into the $\phi(x, y, t)$ field. Branch cuts

begin and end at either branch points or the aperture edge (in effect placing a branch cut just off the aperture at that point). In any path around a branch point, the path will cross the branch cut. The discontinuity encountered at the branch cut is the opposite multiple of 2π than the one encountered in the curl so that the discontinuity at the branch cut counteracts the curl encountered around the branch point and the gradients along the path sum to zero.

Properly placed branch cuts have the effect of making the entire field phase irrotational, allowing phase corrections needed by the field to be implemented by a DM. A continuous facesheet DM implements the phase cut by forming a continuous slope from one pixel to the next over the 2π discontinuity. While non-ideal, this is a significant improvement over simply ignoring the rotational portion of the field [3].

2.11.1 Branch cut impact on DMs. Interestingly, segmented DMs deal with branch cuts very easily. As surface discontinuities are easily implemented, phase corrections are not unwrapped. Leaving the phase wrapped implies that there is a discontinuity leading to every branch point. This discontinuity serves as a branch cut, allowing the DM to correct for the rotational portion of the field.

Continuous facesheet DMs, however, unwrap the phase in order to avoid unnecessary phase discontinuities and make the DM surface as smooth as possible. In this case, branch cuts must be added to create an irrotational correction. Since continuous facesheet DMs cannot implement a surface discontinuity, an area of the DM on either side of the cut cannot achieve the height required for the desired phase correction. As such, a great deal of work has been put into determining the pairing of branch points and placement of branch cuts in order to minimize the impact on system performance. Usually, designs attempt to minimize the length of the cuts while placing the cuts where the least illumination is encountered. Mitigating the effects of branch cuts on system performance in an SRI architecture is discussed in Chapter V.

2.12 Spatial sampling requirements

In order to adequately reflect the optical field and prevent aliasing, the field must be spatially sampled at a minimum of twice the highest spatial frequency. The effect is to

assure that adjacent subapertures in the WFS are within π of each other so that when the field is unwrapped the subaperture phases are properly assigned.

2.13 SRI Development

The SRI has been developed at SOR and tested using their Atmospheric Simulation and Adaptive-optics Laboratory Testbed (ASALT) lab setup. The progress has been substantial and, while still effectively a first-generation sensor, the SRI is an established technology. Initial designs used fiber amplifiers to boost the strength of the reference beams. Current designs are efficient enough that the fiber amplifier (and the accompanying amplification noise) can be eliminated from the design [36]. By removing the optical fiber amplifier, the single mode fiber used as a spatial filter can be as short as 1 cm which further simplifies the design.

2.14 Alternative viewpoint of phase discontinuities

An alternative way of looking at phase discontinuities is to consider the real and imaginary portions of the field $\Re[U(\mathbf{r})]$ and $\Im[U(\mathbf{r})]$. As each is continuous, they form a contour of values in the (x, y) plane. Areas of a contour in the (x, y) plane with positive and negative values are necessarily separated by a line (or possibly an area) where the contour is zero [2]. The intersection of the contours with the zero plane (solutions to $\Re[U(\mathbf{r})] = 0$ or $\Im[U(\mathbf{r})] = 0$) are continuous paths. The paths may be closed within the aperture or extend from edge to edge, but they are continuous. If the contour only touches the zero plane but does not cross it, the intersection is a point (or a line) which is effectively a closed continuous path of zero area.

Figures 24 and 25 were generated by the MATLAB code in Section A.1 of the appendix and show slices through the real and imaginary portions of an example field (shown in Figure 21) where the portions of the field are zero. This field is a result of weak turbulence and has no branch points.

Looked at from this perspective, lines where it is necessary to unwrap the field are apparent. The field is segmented into regions of positive and negative imaginary field components by a line where the imaginary component of the field is zero. The field is

similarly segmented into positive and negative real component regions by a line where the real component of the field is zero. The field should be unwrapped everywhere along lines of zero imaginary component where the real component of the field is negative. Some unwrapping lines may extend from one edge of the aperture to another. Others may start at an edge but terminate where the real portion of the field transitions from positive to negative (this is a branch point as discussed below). Still others may begin and end at branch points.

Branch points are seen as places where $\Re[U(\mathbf{r})] = 0$ lines and $\Im[U(\mathbf{r})] = 0$ lines cross. As the lines cross, both are zero, yielding zero intensity. Furthermore, around the crossing point there are four sections with the real and imaginary portions of the field positive/positive in one section, positive/negative in another, negative/negative in a third and negative/positive in the fourth. Thus the phase of the field sweeps through all four quadrants as the phase is evaluated around a small closed path around the intersection point, and an integral of phase differentials around that path would be non-zero. Thus the curl of the phase at the intersection is non-zero making the point a branch point. Figure 26 shows the lines where the real and imaginary portions of the field used for Figures 24 and 25 are zero overlaid on a single plot. Although the contour lines touch in places, as there are no branch points, they never cross.

Increasing the strength of the turbulence induces branch points. Branch points begin to occur when the turbulence becomes strong enough that the Rytov number is approximately 0.2 or greater [3]. Figure 27 shows a plot of the two contour lines for a field generated by turbulence with Rytov variance 0.1 where there are no branch points. Figure 28 shows a plot of the two contour lines for a field corrupted by turbulence with Rytov variance of 0.4. In this case branch points are present.

The positions of the branch points are determined independently by the MATLAB code `bpfinder()` included in Section A.2 of the appendix. In some cases it appears that there are branch points where the contour slice lines only touch but do not cross. This is a case where the contour slice lines actually do slightly overlap, but the discrete nature of the sampling of the optical field makes it appear that they are instead lying on top of each other.

The contour slice mapping gives unique insight into how branch points should be paired. Consider the case where a $\Re[U(\mathbf{r})] = 0$ line and $\Im[U(\mathbf{r})] = 0$ line are both closed paths (so that they do not extend to the edge of the aperture.) Then the number of intersections (if any) between those two paths is necessarily even. This shows how to pair up the branch points generated by the crossings of the contour slice lines. A branch point should be paired with an adjacent crossing of the same $\Re[U(\mathbf{r})] = 0$ and $\Im[U(\mathbf{r})] = 0$ lines. Consider the simplistic case where $\Re[U(\mathbf{r})] = 0$ lines and $\Im[U(\mathbf{r})] = 0$ lines are overlapping circles. A pair of branch points are easily identified as the two places where the circles cross. Temporally, $\Re[U(\mathbf{r})]$ and $\Im[U(\mathbf{r})]$ are continuous making the contours undulate smoothly. Thus $\Re[U(\mathbf{r})] = 0$ and $\Im[U(\mathbf{r})] = 0$ lines drift and the points at which their lines cross wander. Consider the case where the real and imaginary contour slices form two circles as in Figures 29, 30 and 31. The circles start off overlapped which forms two branch points. As they move apart the branch points move together. Finally, the two circles no longer overlap but only touch, resulting in the elimination of that pair of branch points.

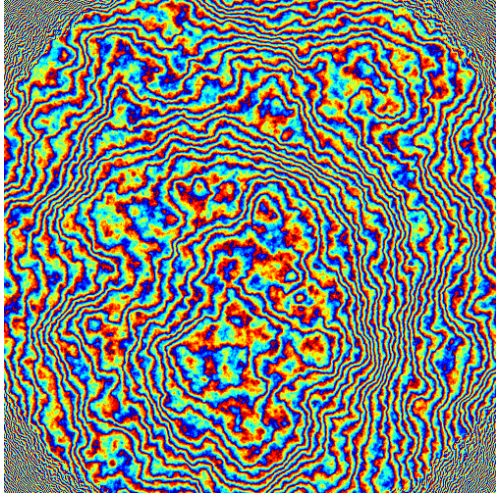
Since branch points occur where $\Re[U(\mathbf{r})] = 0$ and $\Im[U(\mathbf{r})] = 0$ lines cross, both $\Re[U(\mathbf{r})] = 0$ and $\Im[U(\mathbf{r})] = 0$ at branch points. Thus the intensity of the field is necessarily zero at a branch point. While zero intensity is a necessary condition for a branch point, it is not a sufficient condition. This is illustrated by the case where $\Re[U(\mathbf{r})] = 0$ and $\Im[U(\mathbf{r})] = 0$ lines touch but do not cross. At the point they touch, both the real and imaginary portions of the field are zero yielding a zero intensity. As they do not cross, however, the point is not be a branch point.

2.15 Chapter conclusions

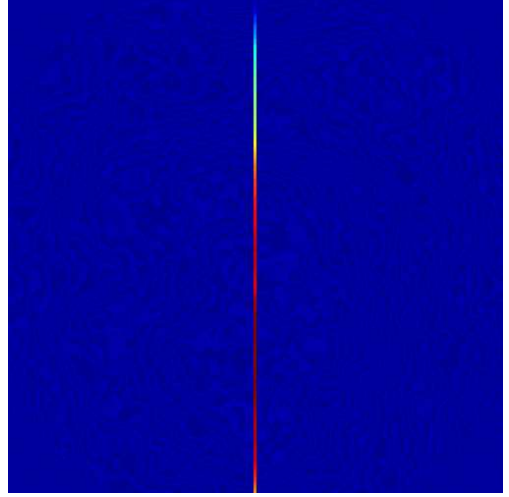
The basics of adaptive optics systems are well documented as are the foundational precepts of strong turbulence. The additional difficulties of operating AO systems under the more challenging conditions of strong turbulence are less well documented. Taking the existing experience in adaptive optics and the understanding of the effects of strong turbulence, this research develops an AO system which can operate effectively under strong turbulence conditions.

In particular, this research utilizes the understanding of conventional AO system designs, the basis of stochastic modeling and estimation, and the basic building blocks of a linear Kalman filter to argue that AO systems can be looked at as fixed gain Kalman filters which estimate the phase of atmospherically aberrated light. The work goes on to consider the architecture of conventional systems and demonstrate why their performance degrades under strong turbulence conditions. Finally the nature of branch points and the localized pairing of branch points are building blocks upon which an unwrapper designed to be effective under strong turbulence conditions is developed.

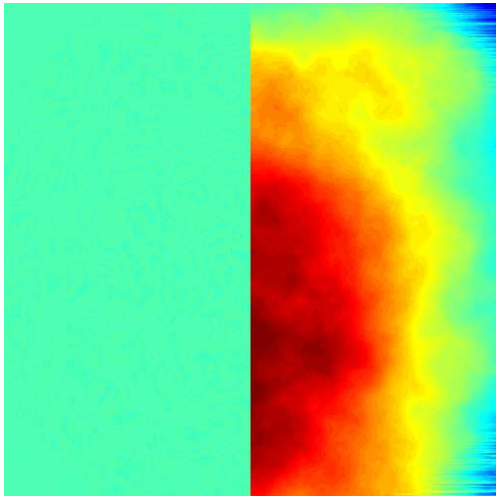
A – Wrapped



B – Center column unwrapped



C – Half unwrapped



D – Fully unwrapped

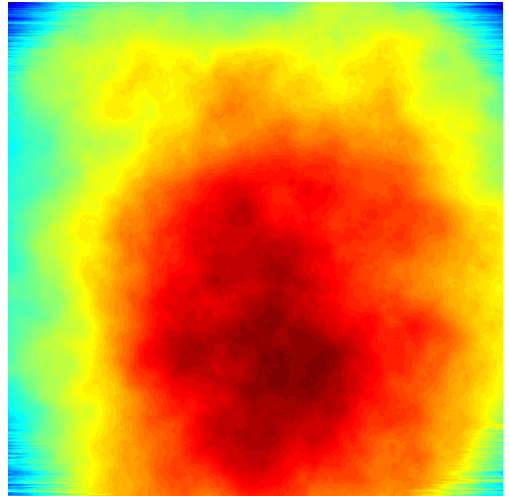


Figure 21: An unwrapping process. First a single column is unwrapped in the center of the field. This avoids spurious data in the corners of the field due to the circular mask of the aperture. Next the field is unwrapped from the center column to the outsides. The right half is unwrapped first, then the array is flipped and the remaining half is unwrapped.

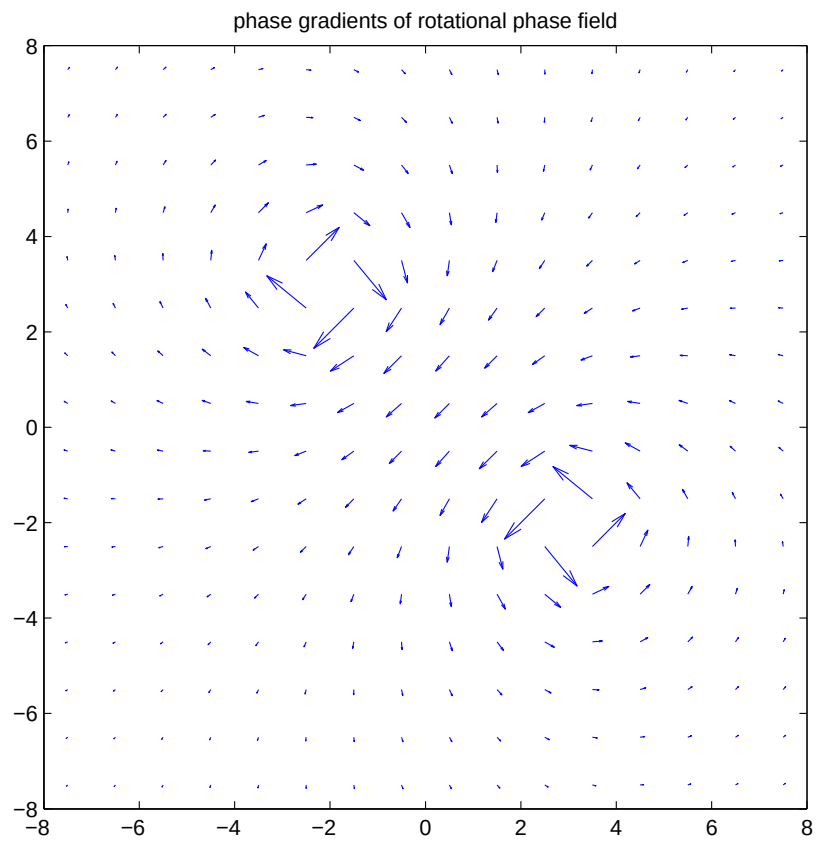


Figure 22: Phase gradients for a field containing branch points.

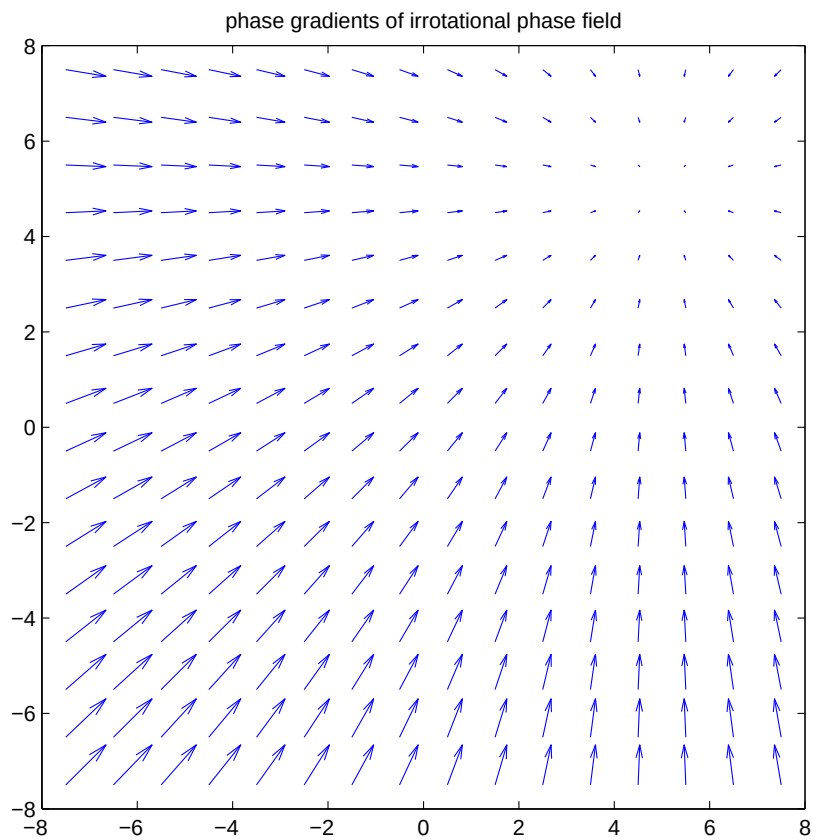


Figure 23: Phase gradients for a field without branch points.

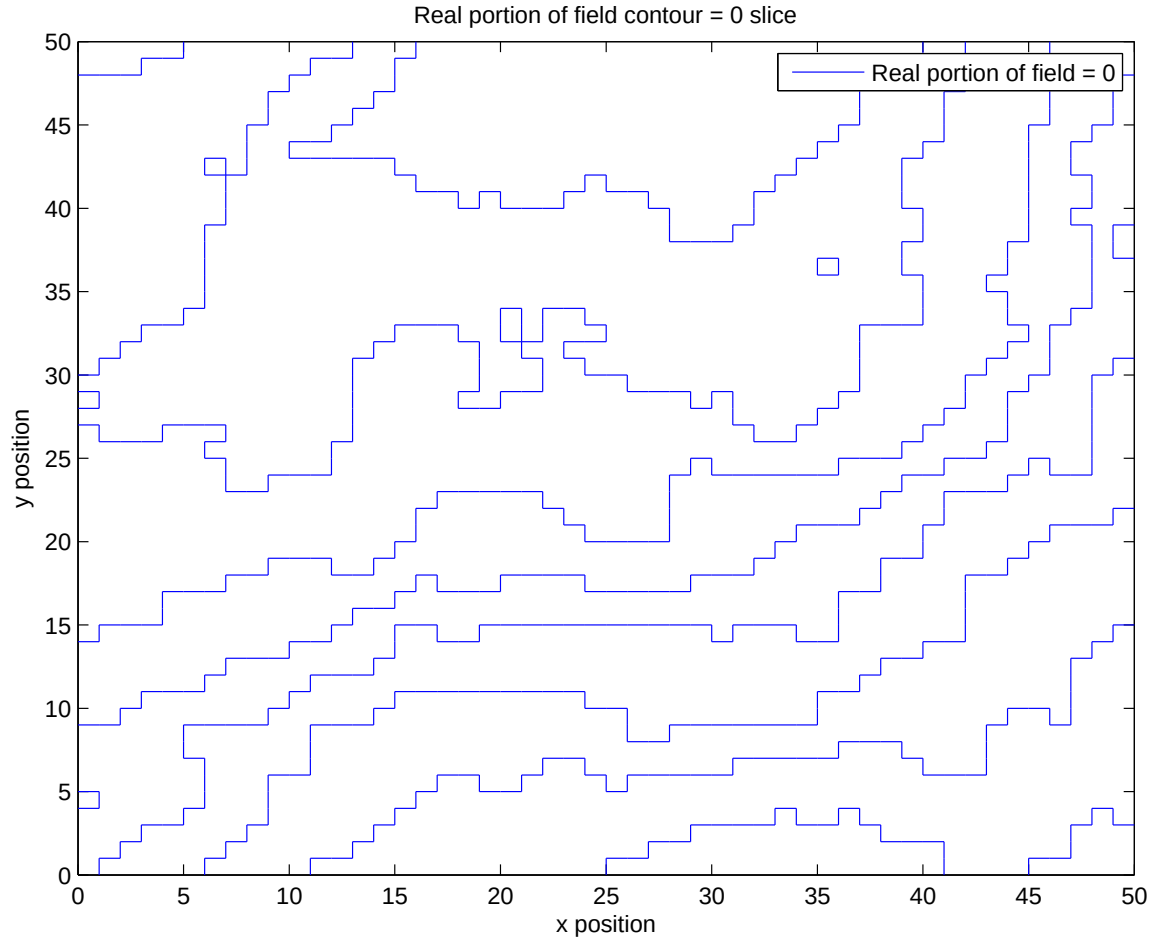


Figure 24: A slice through the real portion of an optical field where the real portion of the field equals zero. The optical field is a 50×50 section of an optical field generated by Wavetrain[®] of an idealized point source through weak atmospheric turbulence.

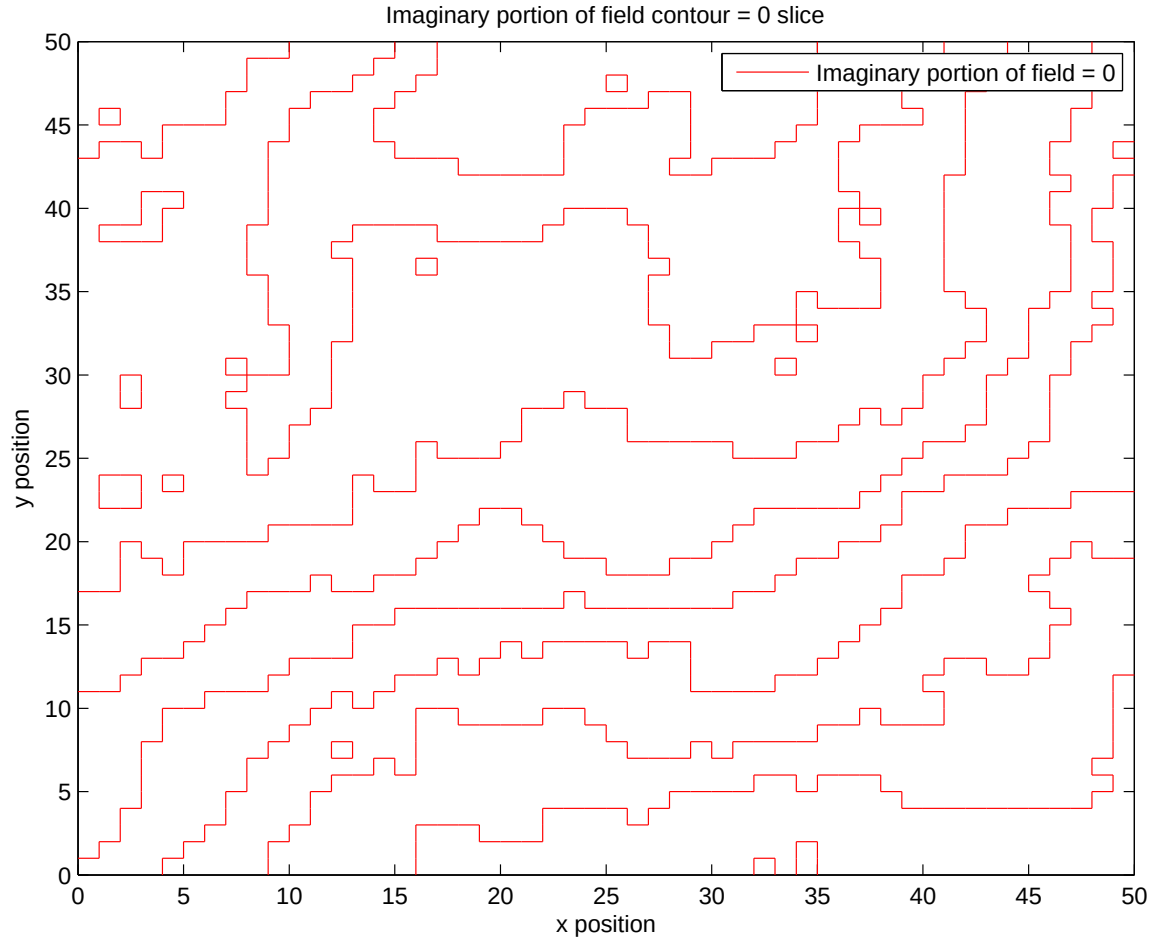


Figure 25: A slice through the imaginary portion of an optical field where the imaginary portion of the field equals zero. The optical field is the same 50×50 section of an optical field depicted by Figure 24.

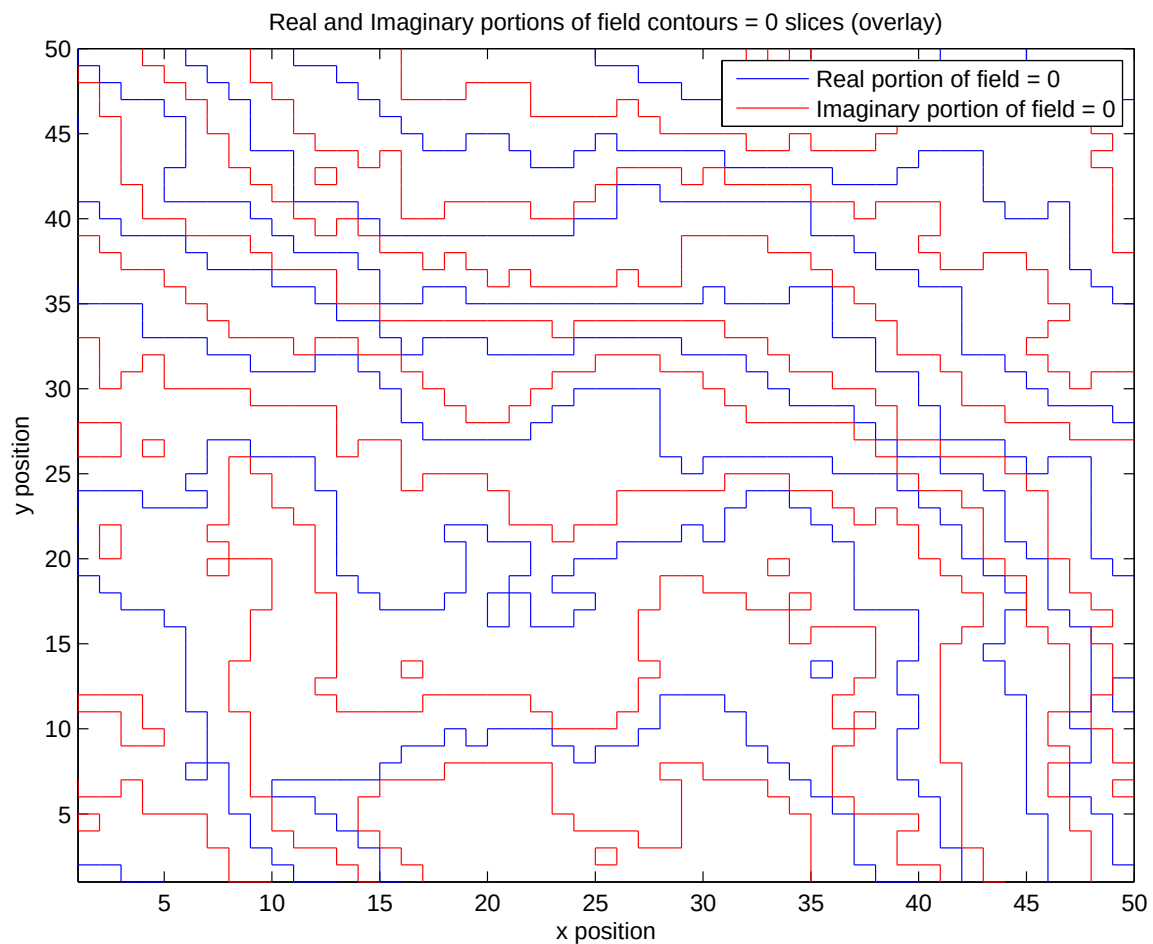


Figure 26: An overlay plot of the real and imaginary contours slices in Figures 24 and 25. Note the areas where two similar contours separated by an opposite contour (blue-red-blue) are only two pixels apart. This indicates that the field is at best critically sampled.

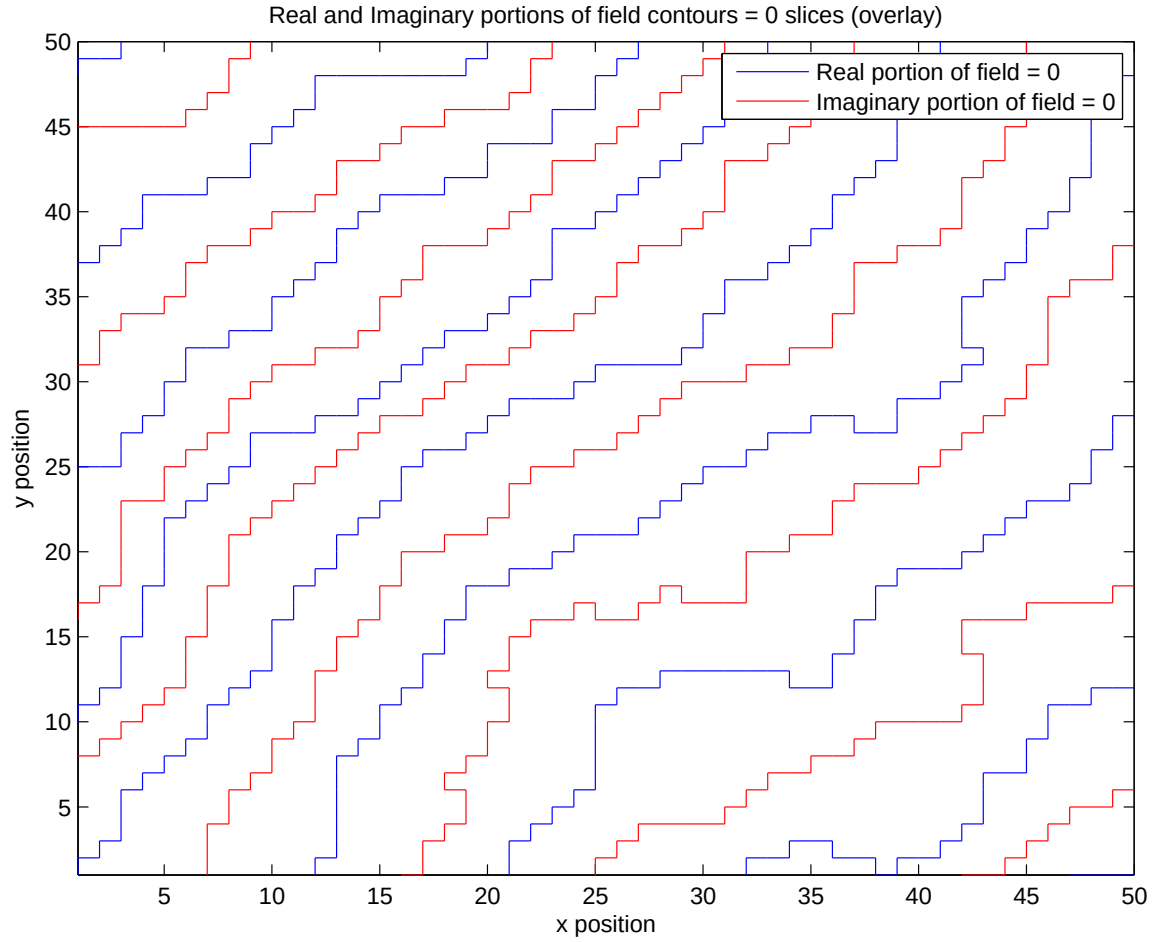


Figure 27: A sliced contour plot for a field corrupted by atmospheric turbulence with Rytov number 0.1. Note the increased spacing between contour lines as compare with Figure 26 indicating a more comfortable spatial sampling rate.

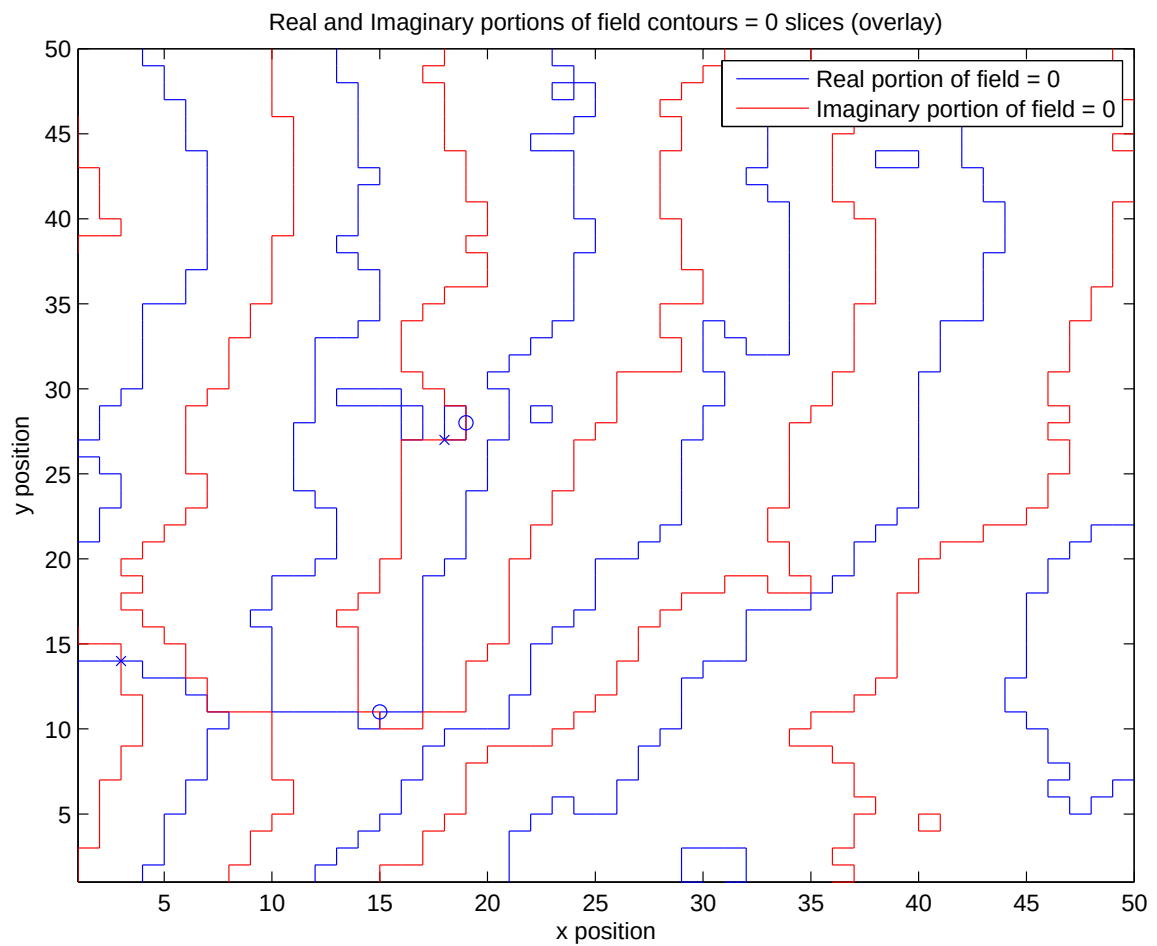


Figure 28: A sliced contour plot for a field corrupted by atmospheric turbulence with Rytov number 0.4. Note the presence of branch points in the field.

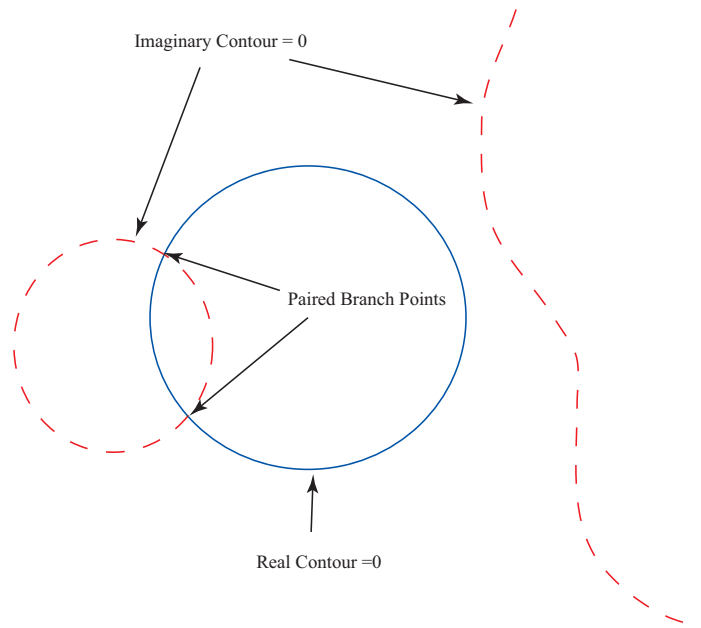


Figure 29: Frame 1 - A depiction of real and imaginary contour slices and the branch point produced where they intersect.

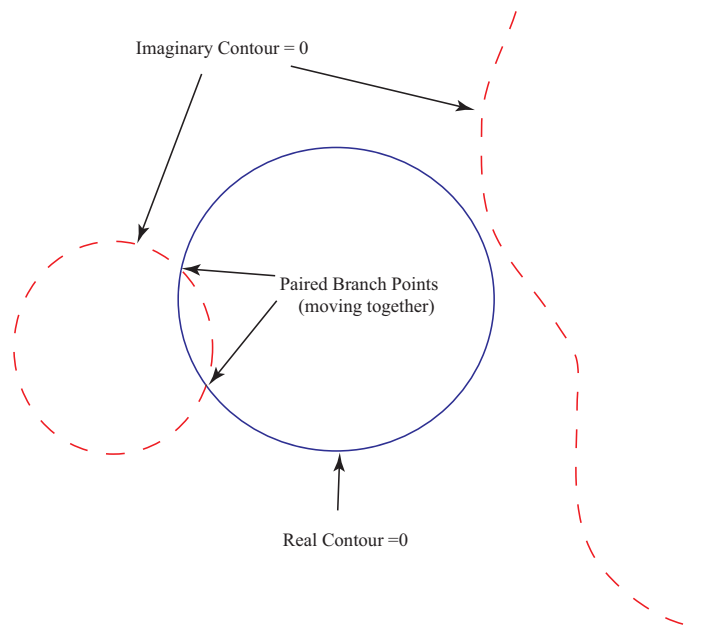


Figure 30: Frame 2 - Note that the branch points are still paired similar to frame 1 in Figure 29, but they are moving closer together as the circles separate.

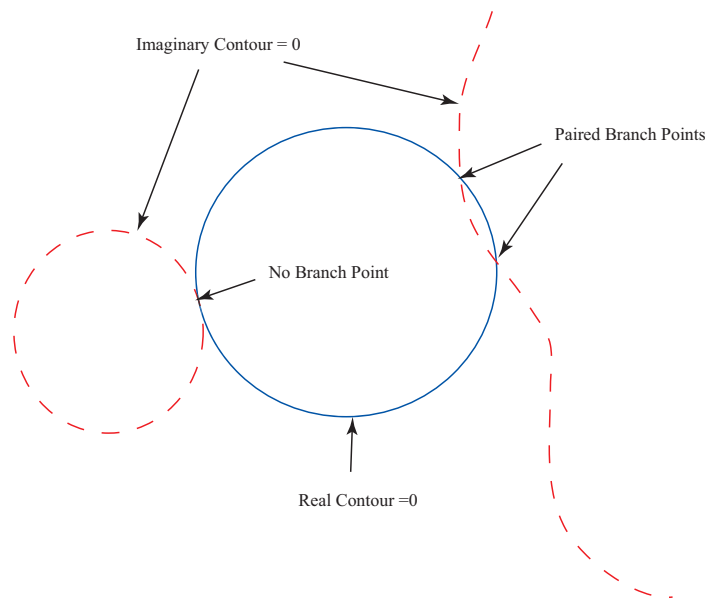


Figure 31: Frame 3 - The contours have moved to the point where the lines of the circles no longer cross but only touch. Thus the circles do not cause a branch point, but new branch points are created by the real circle and an imaginary portion = 0 line that spans the aperture from edge to edge.

III. Adaptive optics, the optical Kalman filter

3.1 Introduction

The intent of this chapter is to document what I believe is the proper treatment of AO from a controls perspective. More specifically, atmospheric turbulence should be considered a stochastic system which induces aberrations into an optical field. An AO system creates an estimate of those aberrations and applies them with a wavefront correction device such as a DM.

3.2 Estimation and correction versus control

Historically, AO systems have been treated as controllers in that they are designed to minimize or control the effects of turbulence. A block diagram of a conventional controller utilizing a DM as a control element is shown in Figure 32. This controller uses a leaky integrator where $u(t_i) = \alpha c(t_{i-1}) + \beta \phi(t_i)$ where $\alpha \approx 0.99$ is one minus the “leak” of the system and $\beta \approx 0.4$ is the proportion of the error signal being integrated. The error signal is the phase output from the WFS and the output of the controller is the phases to be corrected by the DM.

A stochastic model of the system alone is depicted in Figure 33. Here the effects of atmospheric turbulence are the result of a vector of white-Gaussian noise inputs driving an integrator with a (perhaps non-linear) feedback $f(x_1(t), t)$. The result is the state-space $x_1(t)$ which completely defines the optical field prior to reflecting from the DM. After the DM, the state space will be $x_2(t)$ which will be equivalent to $x_1(t)$ except shifted by the phase applied to the DM.

As discussed in Chapter II, optimal stochastic control of such a system is accomplished by measuring the outputs of the system, applying a Kalman filter to the measurements and then using a proportional integrator (PI) controller to achieve the desired type-1 controller property. This is depicted in Figure 34 where ϕ_{Corr} is the phase of the optical field *after* being corrected by the DM.

Expanding the Kalman filter into the block diagram of a linear or at least linearizable Kalman filter is depicted in Figure 35 along with the expansion of a PI Controller into its block diagram form.

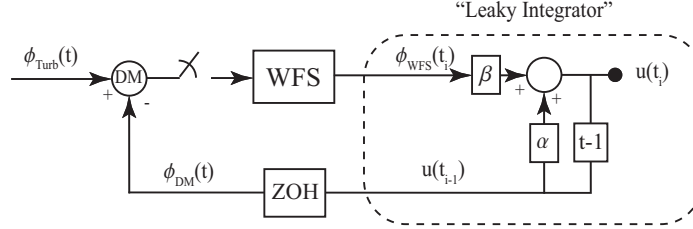


Figure 32: Block diagram of a leaky integrator

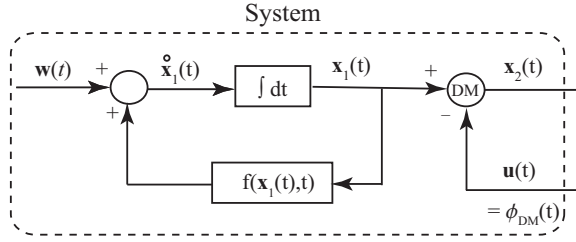


Figure 33: Stochastic model of turbulence with DM

The filter needs to account for the control inputs of the system. This is accomplished by adding the phase applied to the DM to the signal at the beginning of the filter. A more conventional control structure where the control input to the system was able to be applied before the integral in the stochastic system model would instead add the integral of the input with the form $\int_{t_{i-1}}^{t_i} \Phi(t_i, t_{i-1}) \hat{x}(t_{i-1}^+) dt$ [28].

Interestingly, isolating the right hand portion of this block diagram as in Figure 36 highlights that in this case the output of the PI controller is simply an estimate of $\hat{\phi}_{Turb}$ from the middle of the Kalman filter before re-applying the effects of the DM in order to create $\hat{\phi}_{WFS}$.

Thus, such a system would effectively take an estimate of an estimate and is labeled $\hat{\hat{\phi}}_{Turb}(t_i)$. More specifically, it would take an estimate of a Kalman estimate. As a Kalman

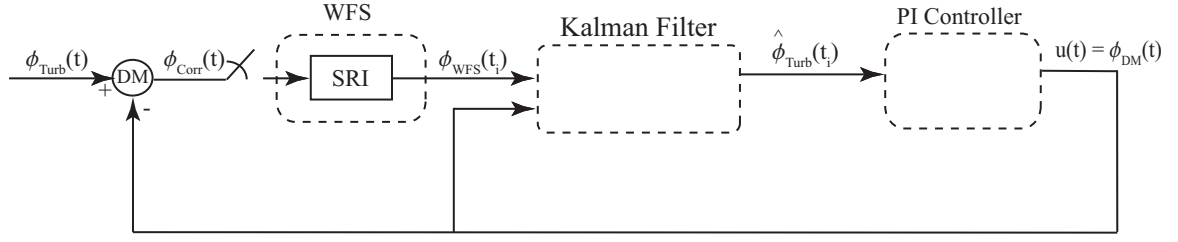


Figure 34: Optimal control structure block diagram

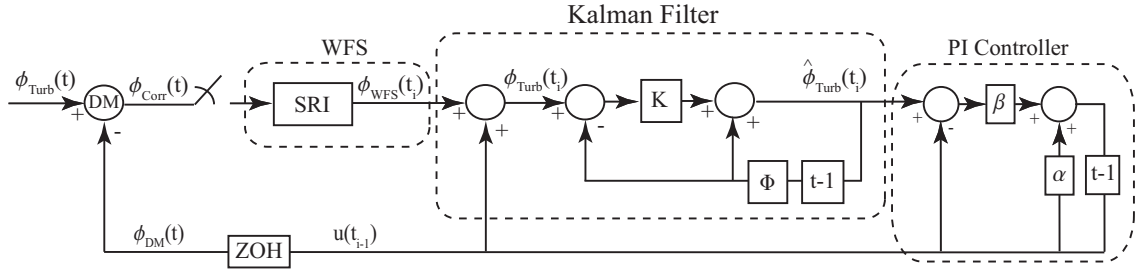


Figure 35: Optimal control structure

filter is an optimal estimator, taking a second estimate is patently unnecessary and potentially degrading to the system.

The essence of the problem is in the placement of the controlling element in the system. Being situated *after* the integrating operation in the stochastic model instead of *before* has the effect of *correcting* the system instead of *controlling* the system. As this differentiation is not widely used in control literature, it is defined here for the purposes of this research. A controlling input enters the stochastic model at the summing junction before the integral operation while a correcting input enters the model at a summing junction after the integral operation. A controlling and correcting system are shown in Figure 37.

In the specific case of AO, the intent is to flatten the wavefront as much as possible. In control terms, the phase of the optical field should be regulated to zero (or a constant piston) after reflection from the DM. The DM deforms in such a way that the phase of the optical field *after* reflecting off of the DM is effectively the phase of the optical field

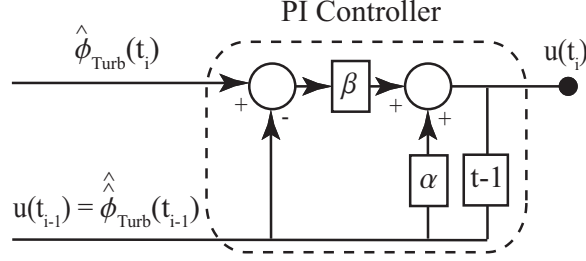


Figure 36: Portion of optimal control structure

before encountering the DM minus the input phase applied to the DM. This is why DMs are symbolized as summing junctions with a negative sign by the input to the DM in the various figures depicting AO systems. Since the DM effectively corrects the optical field by subtracting whatever input phase it is given, the natural choice of control input to the DM is the best estimate of the phase of the optical field striking the DM.

From this perspective, an optimal optical corrector is best depicted as in Figure 38. This isolates the system as the aberrated optical field before it strikes the DM as shown in Figure 39. The WFS measures the phase of the optical field and the Kalman filter creates an optimal estimate of the phase of the field from the WFS measurements. The important point here is that it recognizes the system as uncontrollable. This makes sense in that we have no control over the turbulence received optical field. All we can do is correct or flatten that field by applying the best possible inputs to the DM.

The problem with the AO system as a turbulence estimator is the lack of feedback from the DM. The inclusion of the DM in an open-loop configuration instead of a closed-loop configuration degrades the performance of the system because even with an optimal estimate of the turbulence, any errors between what is commanded to be on the DM and what is actually on the DM is undetected and therefore uncorrected by closed-loop feedback.

As an alternative, consider Figure 40. In the classic Kalman filter design shown in Figure 40(a), the residual is the *calculated* difference between the measured turbulence $\phi_{WFS}(t_i)$ and the estimate of the turbulence $\hat{\phi}_{Turb}(t_i)$. In the modified Kalman filter shown in Figure 40(b), the residual is the phase of the optical field after being corrected by the DM.

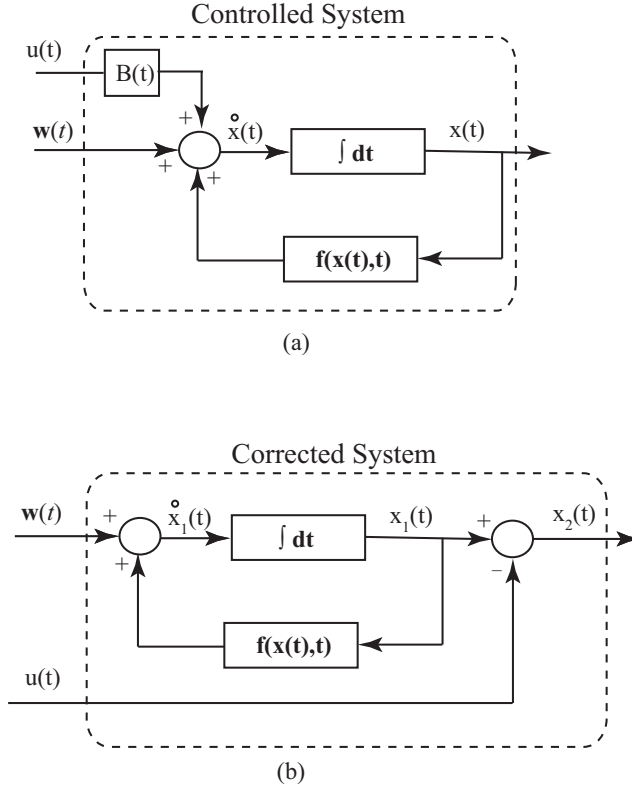


Figure 37: Controlled and Corrected systems

The residual is then *measured* by the WFS. Measuring the residual instead of calculating the residual requires an element capable of performing a difference operation between a received input and an estimation, thus creating the residual to be measured, which is exactly what a DM does in an AO system.

Thus the DM is incorporated into a closed loop. Any errors which are encountered between the DM commands and what is really on the DM are effectively incorporated with the errors associated with the WFS measuring the post-DM optical field.

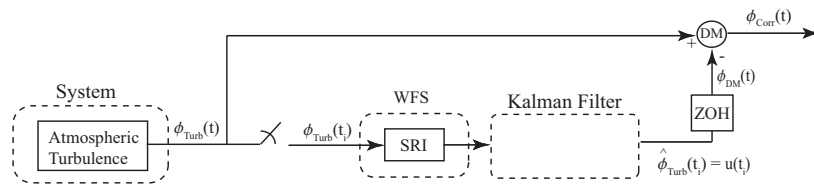


Figure 38: Block diagram of AO estimator

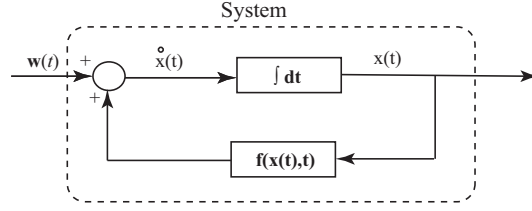


Figure 39: Revised system model

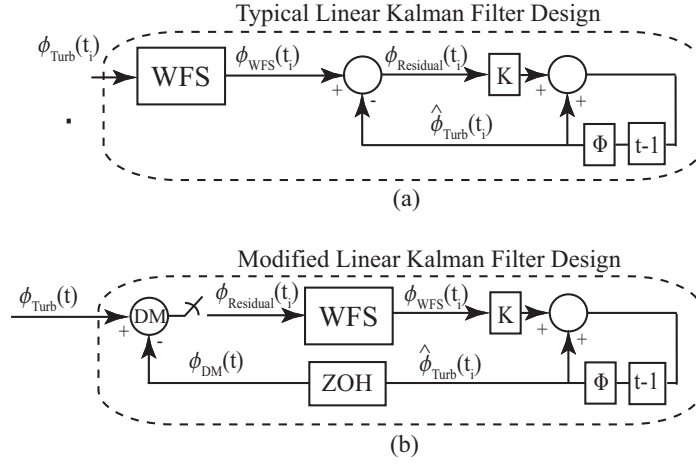


Figure 40: Revised Kalman filter

The final block diagram is depicted in Figure 41 and is recognizable as having the essentially the same form as the leaky integrator already being used in many AO systems, with the Kalman gain K replacing the β and the Kalman state transition matrix Φ replacing the α .

The above discussion may seem a very roundabout way of justifying what is already being done, but there are several advantages in looking at the problem from this perspective. First and foremost, by recognizing an AO system using a PI controller as a Kalman estimator, the design process is justified because a Kalman estimator is an optimal estimator. Rather than reexamine the problem looking at other design architectures, this perspective justifies the architecture already being used and gives a rational for the effectiveness of the current design. Secondly, it gives direction for the improvement of current AO systems using a leaky integrator controller. Recognizing a leaky integrator as a fixed gain Kalman filter, the way to improve performance is to remove assumptions and to refine the system

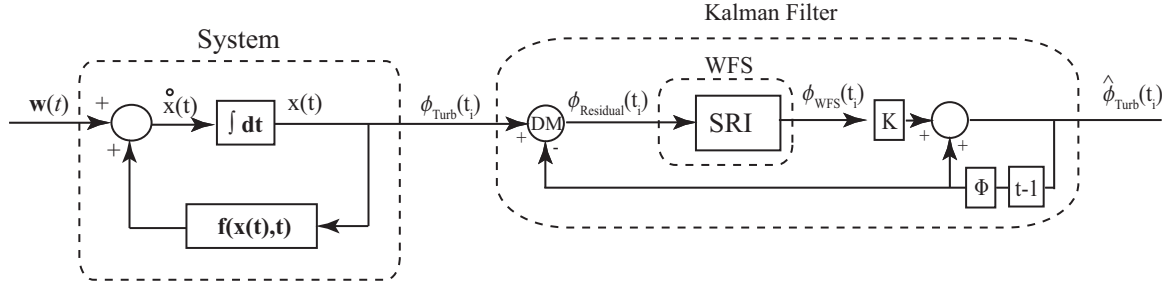


Figure 41: Final system design

model. The biggest assumption in a fixed-gain Kalman filter is that a fixed-gain is adequate. System model inadequacies result from assuming that the system can be modeled by the single state variable $\phi(t_i)$ in a linear model. Adding states and exploring any non-linearities in the system would improve performance.

The bottom line is that there are concrete identifiable ways of increasing the fidelity of the system. Those improvements can be explored for their effectiveness and cost in terms of computational burden and developmental and implementation expense.

3.3 Modeling the system

Effectively modeling the system is key to developing a good controller (or in this case estimator). In a state-based system model, deciding on which states to use in modeling the system is a good place to start.

3.3.1 System states. When choosing the states to model, the things to consider are the system being modeled, what can be sensed about the system, and what states are useful in controlling the system. At any point in time, an optical field can be described by the amplitude and phase of the field. Adding an appropriate number of differential states of amplitude and phase allows modeling of the system dynamics. The number of differential states needed depends on the system and the sampling rate of the system. A system sampled at a much higher rate than the dynamics of the system may be adequately modeled with fewer differential states and may be modeled without any differentials at all.

3.3.1.1 State Observability. What can be sensed about a system is known as the observability of the system in control terminology. Phase, being the typical output of a WFS is the most observable state. Differentials of phase are also observable, but less so than the state directly being sensed. While not typically thought of as observable by an SRI WFS, Section 3.4 discusses a way to derive the amplitudes of the signal and reference, A_S and A_R respectively from SRI data.

3.3.1.2 State controllability. Phase is generally considered to be the only controllable state in a standard AO system because the phase shift caused by the DM is the only control available to an AO system utilizing a single DM. Multi-conjugate systems which employ multiple wavefront correction devices are able to control amplitude, but are not addressed in this research. While uncontrollable by a DM in a single conjugate AO system, having an estimate of the amplitude of the optical field A_S allows improved placement of branch cuts into the less illuminated portions of the field.

3.3.2 System Dynamics. Referring back to Figure 39, the input $w(t)$ is a noise modeling the effects of atmospheric turbulence in aberrating the system. $w(t)$ is a zero-mean white Gaussian noise and has the effect of creating uncertainty between system states at different times. In other words, even if system was known at time t , there will be increasing uncertainty in the estimation of the system at time $t + \tau$, where τ is some time increment greater than zero. In a discretely sampled system, the system dynamics are modeled by $\mathbf{Q_d}(t_i)$, which describes the effects of system noise on the growth of system state uncertainties between samples. In order to examine $\mathbf{Q_d}(t_i)$, a simulation was performed where a series of frames was generated and the changes in phase from the previous frame for each point was recorded. In this simulation, 1000 25×25 subaperture frames were generated with a log-variance of 0.5, $D_{SA}/r_0 = 0.25$, and sample rate $f_S = 44.6f_G$.

Log-variance and D_{SA}/r_0 were set to equal the baseline levels of simulations performed elsewhere during the research. The log-variance is a scintillation index describing the spatial variance of the log-amplitude of the optical field [1]. D_{SA} is the diameter of the subaperture and r_0 is the atmospheric coherence length (sometimes known as Fried's parameter) so that D_{SA}/r_0 describes the ratio of subaperture size to the size of the atmospheric turbulence.

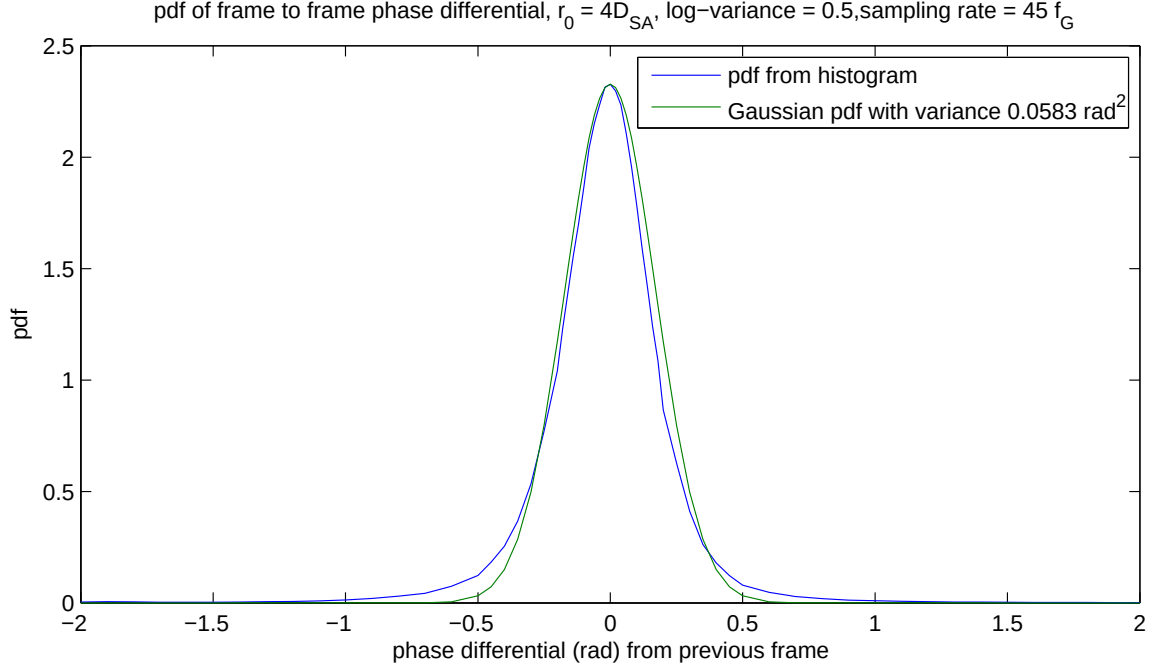


Figure 42: Histogram of subaperture phase changes from frame to frame

This in effect describes the resolution with which the optical field is sensed. The sampling rate was set to a nominal AO sample rate.

It should be pointed out that because of the modulo- 2π nature of the measured phase, the changes recorded were restricted to $(-\pi, \pi]$. Thus a point which had a phase of 0.9π in one frame and then -0.9π the next would be interpreted as having changed $+0.2\pi$ instead of changing -1.8π . A histogram of the frame-to-frame changes is shown in Figure 42 and describes the pdf of the phase changes from one frame to the next. The histogram is not perfectly Gaussian but certainly has a zero-mean Gaussian shape with an approximate variance of 0.583.

3.3.3 Measurement noise. Determining the uncertainty characteristics or noise pdf of the output of an SRI is difficult. Consider Figure 43 where the evaluation of the phase of a single sub-aperture in the wavefront is graphically depicted. Here the measurements are the square of the distance between reference points and the point described by the phase and amplitude of the field in the complex plane. The phase ϕ_{WFS} can then be determined

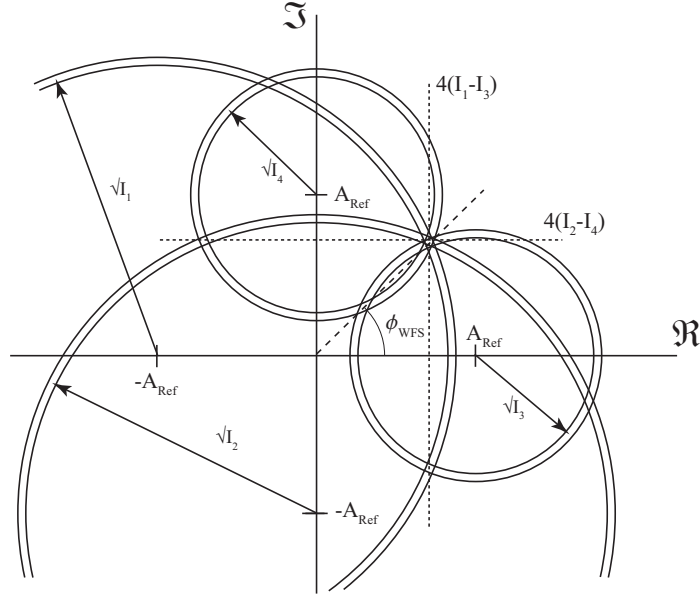


Figure 43: Graphical depiction of SRI phase estimation

as the arctangent of the ratio between the differences, $\phi = \tan^{-1}[(I_2 - I_4)/(I_1 - I_3)]$ as previously described in Equation (12).

Individual interferograms are well modeled as a measurement corrupted by read and shot noises. The read noise is an artifact of the camera and is assumed to have a zero-mean Gaussian pdf [23]. The shot noise is well-modeled as a Poisson process [13]. For the purposes of this research, the signal is assumed to be strong enough that the shot noise can be modeled as a zero-mean Gaussian pdf with a SNR of the square root of intensity of the signal. The uncertainty of interferogram measurements I_1 through I_4 is depicted in Figure 43 by a double line, indicating that the actual values likely lie somewhere between the two lines. The impact of this uncertainty in the interferogram measurements on the computed phase ϕ_{WFS} is difficult to determine however in that the arctangent function used to convert the interferograms into phase is highly non-linear.

Rather than determine a mathematical uncertainty of ϕ_{WFS} based on the interferogram uncertainties, a Monte Carlo analysis was performed in a simulation of an SRI WFS. Here a system is defined with a known signal amplitude and phase, reference signal amplitude as well as interferogram read noises generated by the MATLAB Gaussian number generator. one million realizations of this system were executed in a MATLAB simulation.

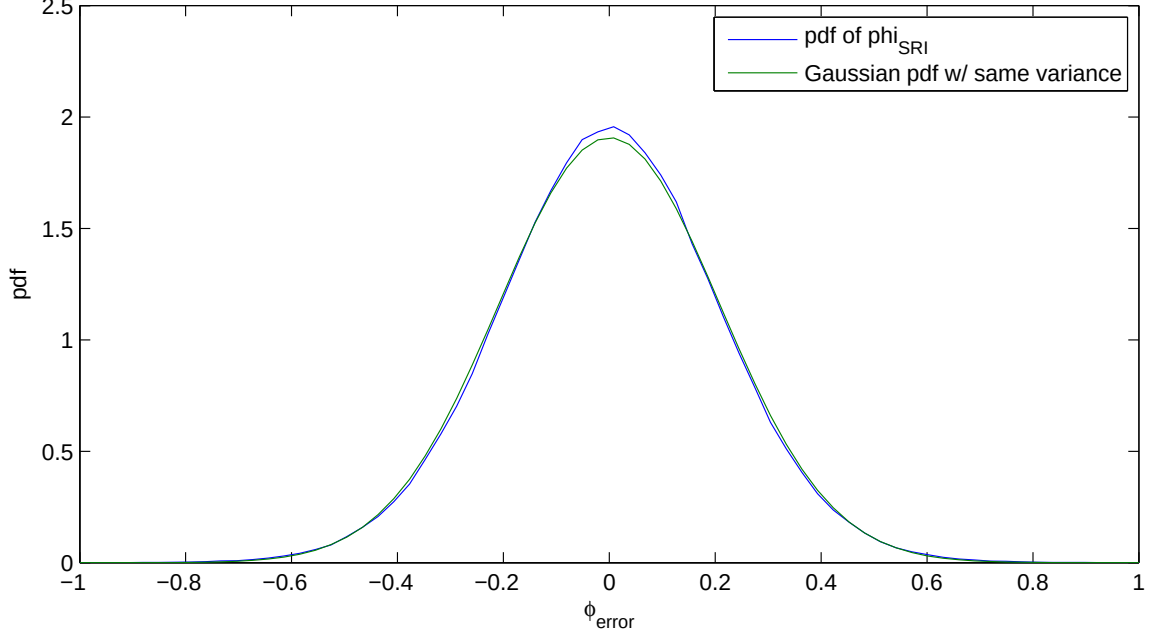


Figure 44: Empirical pdf of ϕ_{SRI} with moderate noise

Results were recorded as the difference between the phase output of the simulation and actual ϕ and put into a histogram to generate an empirical pdf of the SRI phase output error. The system was set up with the phase $\phi = \pi/4$, signal and reference amplitudes $A_S = A_R = 1$, read noise variance $\sigma_{\text{Read}}^2 = 0.3$, and a shot noise photons per intensity unit of 60. Parameters are chosen to generate a reasonably noisy system. This simulation is read noise dominated and has SNRs for the lower intensity interferograms of 1.9 while the higher intensity interferograms have SNRs of 9.6. The result is Gaussian in nature and shown in Figure 45 where it is compared against a Gaussian distribution with similar variance.

Only under the strongest noise conditions does the pdf become less Gaussian. Figure 45 shows the results from increasing the read noise to one where the SNRs vary between 0.58 and 3.2 for the interferograms. In this case the pdf is still Gaussian-like, but narrower with higher tails than a Gaussian of similar variance would have.

The code to generate these plots is in Section A.3 of the appendix. From these plots it is reasonable to treat the output from an SRI ϕ_{SRI} as a measurement of ϕ corrupted by a zero-mean Gaussian noise.

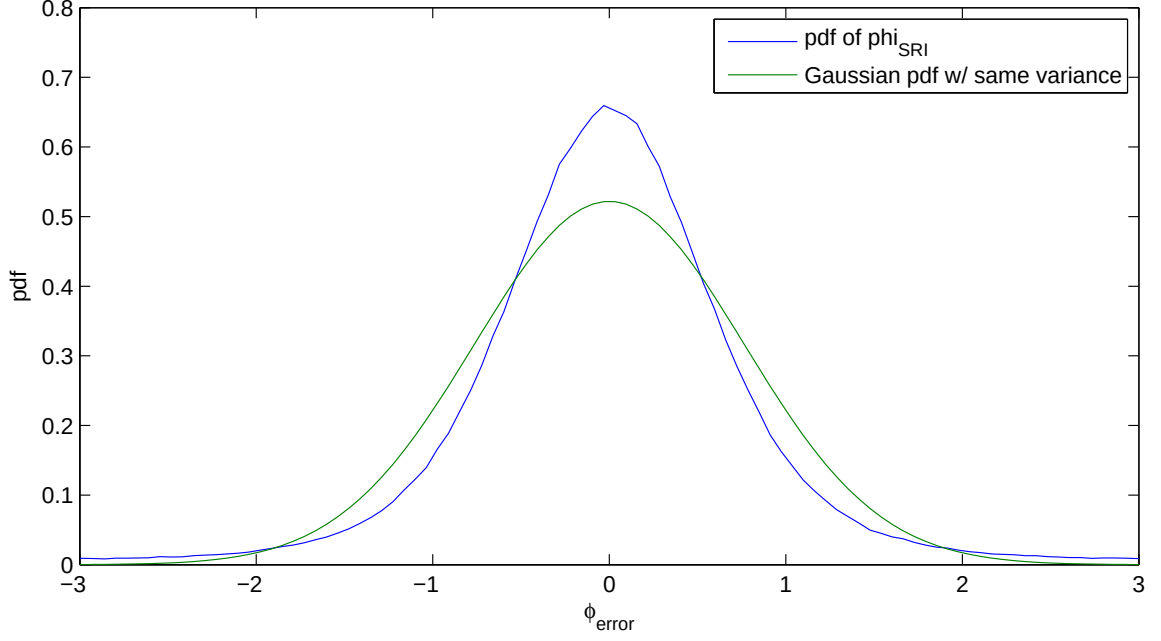


Figure 45: Empirical pdf of ϕ_{SRI} with high noise

The variance of the measurement noise may be impacted by the conditions of the system. Conditions to be considered are the signal and reference amplitudes A_S and A_R , signal phase angle ϕ , strength of the camera read noise σ_{Read} , and photons per intensity unit Q . Empirical pdfs for measurement error were generated while parameters were varied one at a time and the variance of the pdfs was recorded.

The first variable to be examined is ϕ . As such, ϕ is varied from 0 to 2π and the variance of the difference between ϕ_{Est} and ϕ is plotted in Figure 46. See Section A.4 in the appendix. In Figure 46, conditions are similar to Figure 44 except for ϕ varying between 0 and 2π .

The flatness of Figure 46 was consistent for various noise conditions in both shot and read noise dominated systems. From this it can be concluded that the variability of ϕ_{SRI} is not dependent on ϕ .

Considering the other variables, the noise of the system is a summation of the shot and read noise and the amplitudes A_S and A_R are strongly related to the signal strength. As such the variance of the system is dependent on the SNR which in this case is essentially the ratio of A_R^2 and A_S^2 to the dominant noise source.

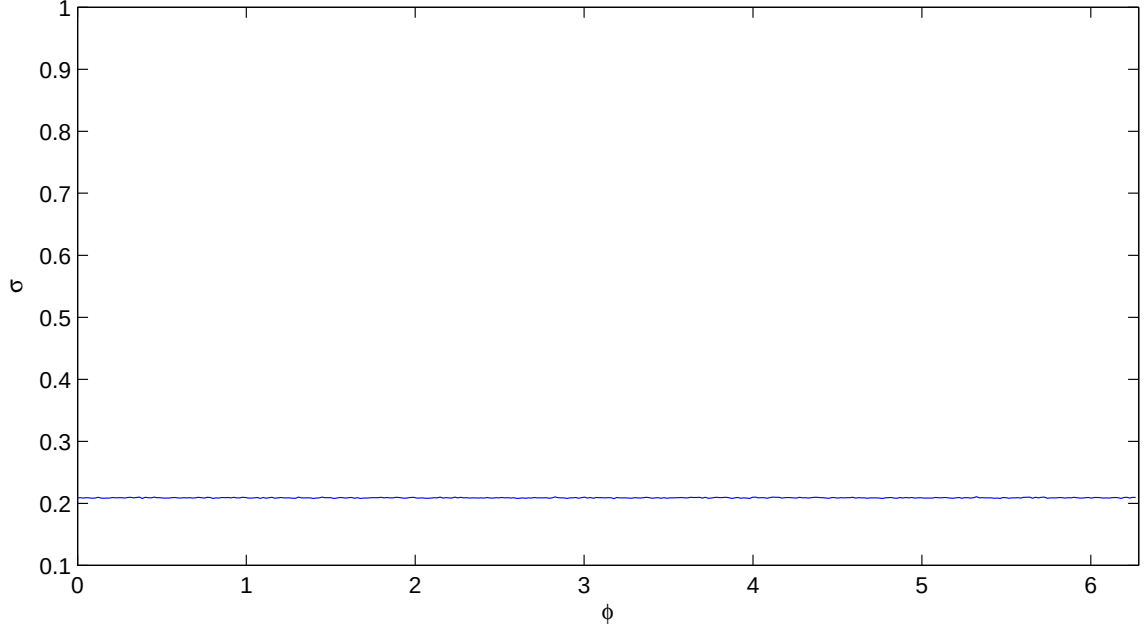


Figure 46: Standard deviation σ vs ϕ

A simulation was performed which varied A_R and A_S under extremely high noise conditions. ‘Truth’ interferograms were developed as

$$I_1 = (A_R + A_S \cos \phi)^2 + (A_S \sin \phi)^2 \quad (28)$$

$$I_2 = (A_S \cos \phi)^2 + (A_R + A_S \sin \phi)^2 \quad (29)$$

$$I_3 = (A_R - A_S \cos \phi)^2 + (A_S \sin \phi)^2 \quad (30)$$

$$I_4 = (A_S \cos \phi)^2 + (A_R - A_S \sin \phi)^2 \quad (31)$$

These interferograms were then corrupted by noise of strength $\sigma_{Noise}^2 = (50000^2 + I_n)^{1/2}$ where the 50,000 accounted for the read noise of the measurement device and the I_n took the shot noise into account. The system was read noise dominated at all points, with A_R and A_S varied from 0 to 260 and 1300 respectively. 40,000 samples were taken at each point and ϕ_{Est} was determined for each sample using the standard arctangent function of Equation 12. ϕ_{Errors} were determined as the difference between ϕ_{Actual} and ϕ_{Est} . The variance of the 40,000 ϕ_{Errors} were plotted in a surface plot as shown in Figure 47.

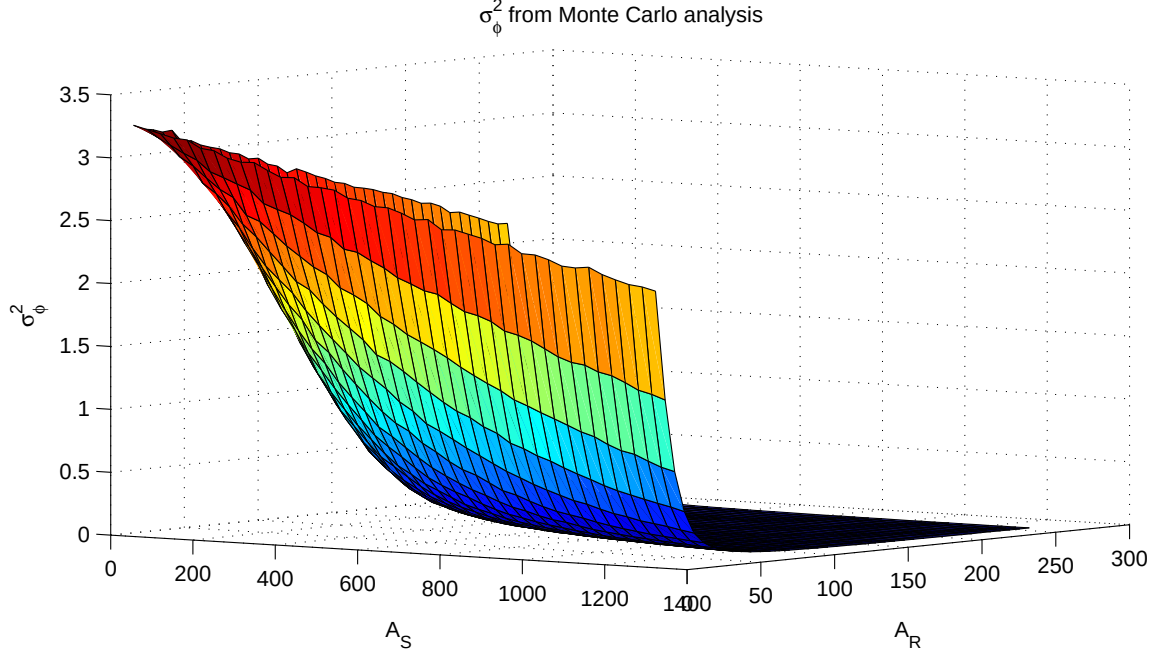


Figure 47: Variance of ϕ_{SRI} determined through Monte Carlo analysis

The curve of the surface suggests a $1/A_R$ and $1/A_S$ dependency and in fact the equation

$$\sigma_{phi}^2 = 3.29e^{\frac{-A_S A_R}{17500}} \quad (32)$$

seems to match the empirically measured variance, as shown in Figure 48. To be clear, Equation (32) does not have a particular theoretical basis but is simply an equation found to match the empirical results. The conclusion is that whenever A_R or A_S is small when compared to the noise of the interferogram measurements, the quality of the measurements is decreased. This makes sense because when the reference amplitude A_R is small, the interferograms are largely the same and it is more difficult to extract meaningful information from the differences. When A_S is small, a small error in the complex domain can lead to a significant error in the phase domain.

The other observation of note is that the estimated variance is really quite small except where A_R or A_S are small. This highlights the relative noise immunity of the SRI. For $(A_S^2/\sigma_{\text{read}}^2)(A_R^2/\sigma_{\text{read}}^2) > 0.37$, the variance of the measured ϕ_{SRI} is less than 0.1 rad^2 . Under closed-loop conditions A_R , while not large, should at least be well above zero. A_S is much more the issue in that under highly scintillated conditions, A_S will vary greatly

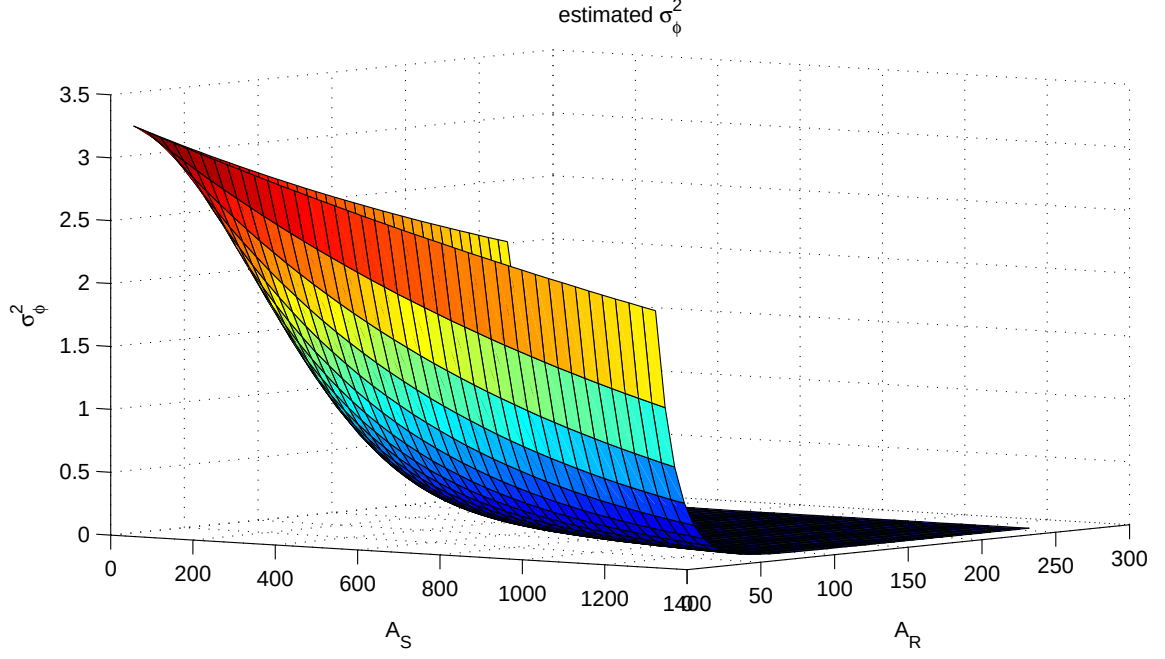


Figure 48: Variance of ϕ_{SRI} determined through Equation (32)

and often be small. Measurements taken in these regions of low intensity will have higher variance but be less critical than measurements of regions of high intensity because of their lower energy content.

Having an idea of the quality of the measurements provided by the SRI WFS allows us to design the Kalman filter estimating the system. Modeling the system and knowing the variance of the WFS measurements allows us to set the Kalman gain of the system. Ideally, the computation of the gain is as depicted by Equation (20) where the variance of the system is dynamically tracked to make optimal use of measurements provided by the sensor. In systems with recurring periodic measurements, however, several simplifications can be made which greatly reduce computational requirements without significantly affecting performance.

The first assumption which has already been discussed is that system dynamics can be encapsulated into a single variable Q_d . The second is that the system can be thought of as in steady state. That is, that the variance of the system immediately after a measurement is essentially the same as the variance of the system immediately after the previous measurement. This is true of systems which are sampled at a higher rate than their system dynamics

and should be true in AO systems sampled above 10 times the Greenwood frequency, which most designers consider a minimum sampling rate. Finally, the state transition matrix is one well-modeled as the identity matrix, which implies that phase does not decay towards a mean between samples. Without prior knowledge, the modulo 2π nature of phase implies that phase should be uniformly distributed between $-\pi$ and π , so that phase will not decay towards a particular value.

Beginning with the variance of the estimate just after the measurement is taken, the following simplifications can be made:

$$P(t_i^+) = P(t_i^-) - K(t_i)H(t_i)P(t_i^-) \quad (33)$$

$$P(t_i^+) = [I - K(t_i)H(t_i)] P(t_i^-) \quad (34)$$

Continuing with the variance of the estimate immediately after the measurement is taken,

$$P(t_i^-) = \Phi(t_i, t_{i-1})P(t_{i-1}^+)\Phi^T(t_i, t_{i-1}) + \int_{t_{i-1}}^{t_i} \Phi(t_i, \tau)G(\tau)Q(\tau)G(\tau)\Phi^T(t_i, \tau)d\tau \quad (35)$$

$$= \Phi(t_i, t_{i-1})P(t_{i-1}^+)\Phi^T(t_i, t_{i-1}) + Q_d(t_i) \quad (36)$$

$$= P(t_{i-1}^+) + Q_d(t_i) \quad (37)$$

$$= [I - K(t_{i-1})H(t_{i-1})] P(t_{i-1}^-) + Q_d(t_i) \quad (38)$$

Setting $P(t_i^-)$ equal to $P(t_{i-1}^-)$ equal to $P(t^-)$, and recognizing that for a system with a single state variable and a one-to-one mapping from measurements to states, $H(t_i) = 1$ we obtain

$$P(t^-) = [I - K(t_{i-1})H(t_{i-1})] P(t^-) + Q_d(t_i) \quad (39)$$

$$= \frac{Q_d(t_i)}{K(t_{i-1})}. \quad (40)$$

Examining the equations for the Kalman gain K ,

$$K(t_i) = P(t_i^-)H^T(t_i)[H(t_i)P(t_i^-)H^T(t_i) + R(t_i)]^{-1} \quad (41)$$

$$= P(t^-)[P(t^-) + R(t_i)]^{-1}. \quad (42)$$

Since P and R will be scalars,

$$K(t_i) = \frac{P(t^-)}{P(t^-) + R(t_i)} \quad (43)$$

$$= \frac{\frac{Q_d(t_i)}{K(t_{i-1})}}{\frac{Q_d(t_i)}{K(t_{i-1})} + R(t_i)} \quad (44)$$

$$= \frac{Q_d(t_i)}{Q_d(t_i) + R(t_i)K(t_{i-1})} \quad (45)$$

now, assuming that $K(t_i)$ does not change significantly from one measurement to the next, $K(t_i) = K(t_{i-1})$, and

$$K(t_i) = \frac{Q_d(t_i)}{Q_d(t_i) + R(t_i)K(t_i)}. \quad (46)$$

Then

$$K(t_i) [Q_d(t_i) + R(t_i)K(t_i)] = Q_d(t_i) \quad (47)$$

$$R(t_i)K(t_i)^2 + Q_d(t_i)K(t_i) - Q_d(t_i) = 0 \quad (48)$$

$$K(t_i) = \frac{-Q_d(t_i) \pm \sqrt{Q_d(t_i)^2 + 4R(t_i)Q_d(t_i)}}{2R(t_i)} \quad (49)$$

and since $K(t_i) > 0$, this becomes

$$K(t_i) = \frac{Q_d(t_i)}{2R(t_i)} (\sqrt{1 + 4R(t_i)/Q_d(t_i)} - 1), \quad (50)$$

and K as a function of $\frac{R}{Q_d}$ is depicted in Figure 49.

Since R is a function of A_R and A_S , and Q_d is constant unless the turbulence characteristics change, the natural conclusion is that $K(A_R, A_S)$ can be determined as $K(A_R, A_S) = Q_d(t_i)/[2R(t_i)]\{ (1 + 4R(t_i)/Q_d(t_i))^{1/2} - 1 \}$. An example of such a relationship is shown in Figure 50.

Implementation, however is not so simplistic. Determining the effects of A_R and A_S on measurement accuracies is difficult and complicated by the fact that when A_R and A_S become small, measurement errors appear to become less Gaussian. In addition, Q_d may

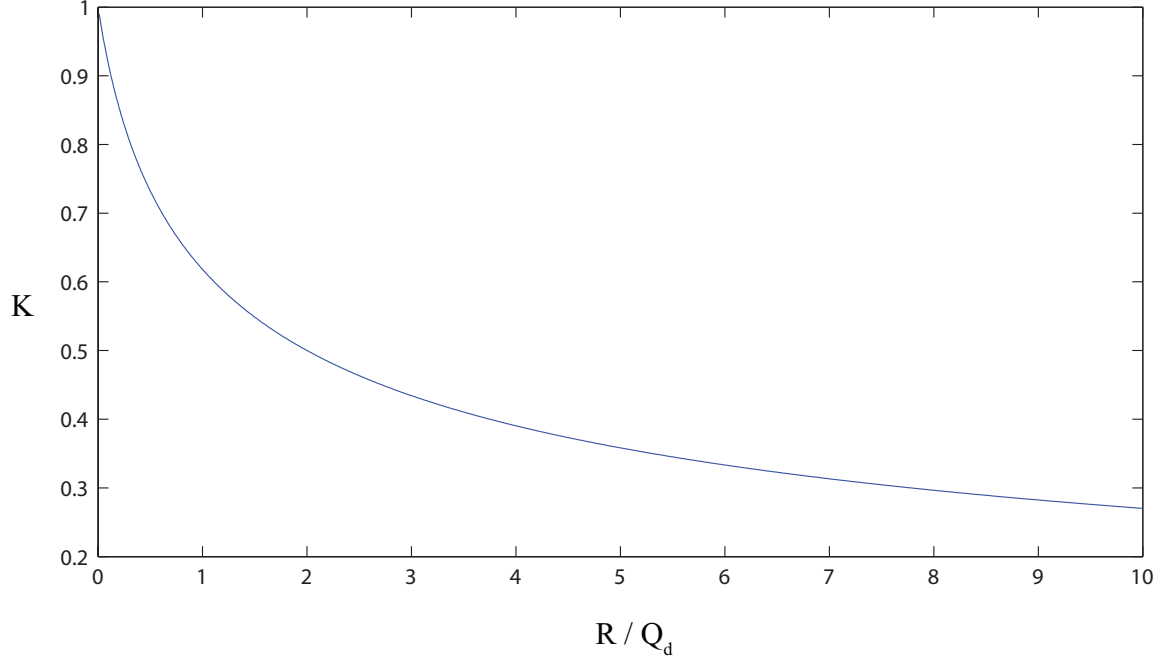


Figure 49: K vs. $\frac{R}{Q_d}$

vary at small A_S values because the complex value of light at that point may jump over and around the origin, causing ϕ to vary more from frame to frame than normal.

A second factor is that areas of high intensity in an optical field have the greatest impact on system performance because that is where the most energy resides. Varying the Kalman gain K at low intensity areas where A_S is small optimizes the system in the low-energy, least-important parts of the field.

A final factor is that in an AO system with fast sampling, the criticality of K is low. If the system is sampled at many times the Greenwood frequency f_G , having the gain K be slightly less than optimal does not affect system performance very much because Q_d is small. The system takes multiple measurements of what is effectively the same system states.

The bottom line is that using a full-blown Kalman filter, or even using steady-state assumptions and simply adjusting K based on an expected variance of the measurements, would be an entire research topic in itself and is beyond the scope of this research. Fixed gain

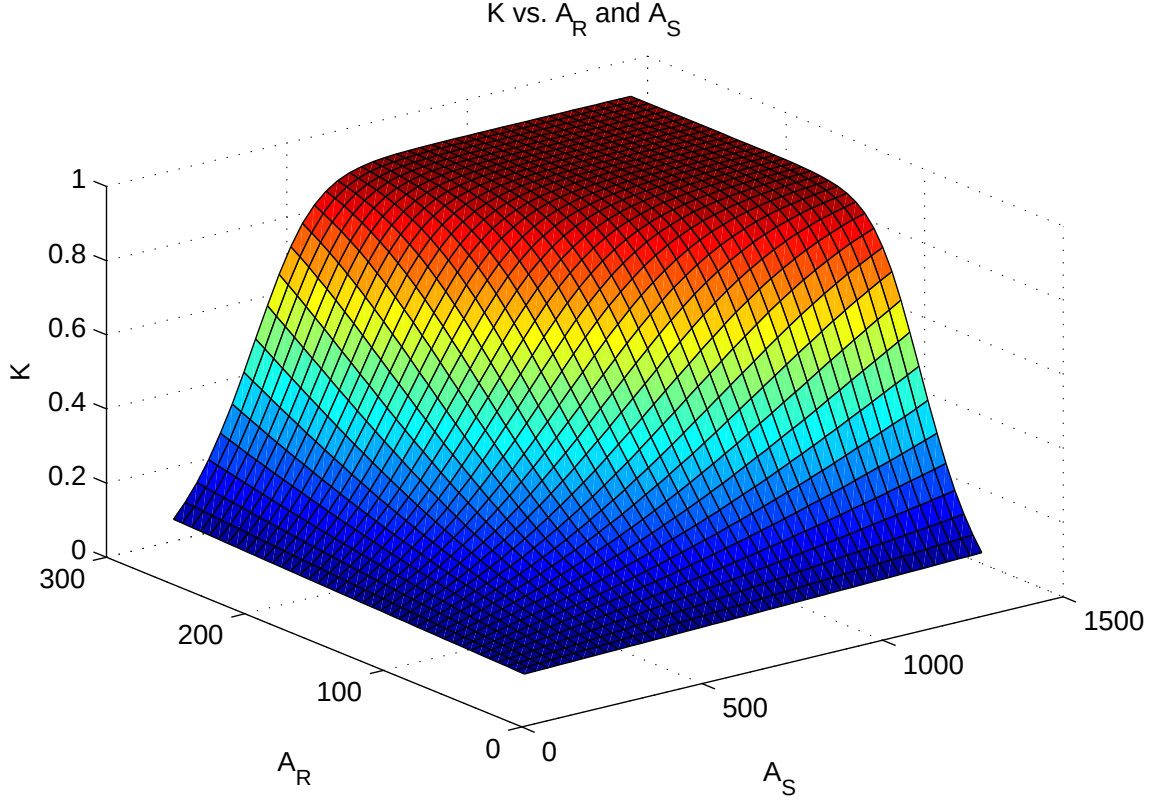


Figure 50: K vs. A_R and A_S

proportional integrators work extremely well for highly sampled systems with reasonable SNRs, even in strong turbulence, and is used in this project.

3.4 Deriving wavefront amplitude from SRI output

As mentioned previously, the phase of a wavefront can be determined as the arctangent of the ratio of interferogram differences.

$$\phi = \tan^{-1} \left(\frac{I_1 - I_3}{I_2 - I_4} \right) \quad (51)$$

Generally knowing the phase is sufficient, but knowing the amplitude would be potentially advantageous, particularly in highly scintillated fields where the amplitude varies significantly. Towards that end, it would be useful to be able to determine wavefront amplitude from SRI interferogram measurements.

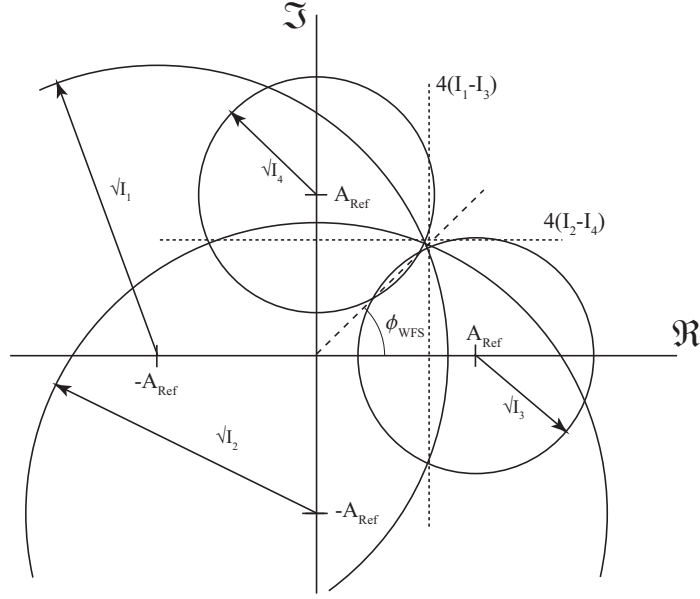


Figure 51: Graphical depiction of SRI phase estimation

Isolated to a single point, as depicted in Figure 51, the interferograms of an SRI describe a system as

$$\begin{aligned}
 I_1 &= (x + A_R)^2 + y^2 \\
 I_2 &= x^2 + (y + A_R^2) \\
 I_3 &= (x - A_R)^2 + y^2 \\
 I_4 &= x^2 + (y - A_R)^2,
 \end{aligned}$$

where x is the real component of the field, y is the imaginary component of the field, and A_R the amplitude of the reference. Essentially, this is an over-determined system with four equations and the three unknowns of x , y , and A_R . Conversely, the unknowns can be treated as A_S and ϕ where $x = A_S \cos(\phi)$ and $y = A_S \sin(\phi)$. This is convenient because as previously mentioned, ϕ is already solved by the arctangent function. Useful manipulations of the aforementioned interferogram equations are

$$\begin{aligned}
 \frac{I_2 - I_4}{I_1 - I_3} &= \frac{4y}{4x} \\
 &= \frac{y}{x},
 \end{aligned}$$

which is used by the arctangent function to determine ϕ ,

$$(I_1 - I_3)^2 + (I_2 - I_4)^2 = (4xA_R)^2 + (4yA_R)^2 \quad (52)$$

$$= 16A_R^2(x^2 + y^2) \quad (53)$$

$$= 16A_R^2A_S^2 \quad (54)$$

and

$$I_1 + I_2 + I_3 + I_4 = 4(x^2 + y^2 + A_R^2) \quad (55)$$

$$= 4(A_S^2 + A_R^2). \quad (56)$$

Now, rearranging Equation (55), we write

$$-4(A_S^2 + A_R^2) + I_1 + I_2 + I_3 + I_4 = 0.$$

Multiplying through by A_R^2 and dividing by -4 gives

$$A_R^4 + A_R^2A_S^2 - \frac{I_1 + I_2 + I_3 + I_4}{4}A_R^2 = 0.$$

Rearranging again and using Equation (52) to substitute for $A_R^2A_S^2$ leaves

$$\begin{aligned} A_R^4 - \frac{I_1 + I_2 + I_3 + I_4}{4}A_R^2 + A_R^2A_S^2 &= 0 \\ A_R^4 - \frac{I_1 + I_2 + I_3 + I_4}{4}A_R^2 + \frac{(I_1 - I_3)^2 + (I_2 - I_4)^2}{16} &= 0 \end{aligned}$$

which is a quadratic of A_R^2 and can be solved as

$$A_R^2 = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

where $A = 1$, $B = -(I_1 + I_2 + I_3 + I_4)/4$, and $C = [(I_1 - I_3)^2 + (I_2 - I_4)^2]/16$. This provides two answers and it can in fact be shown that one answer is A_R^2 while the other answer is A_S^2 . The dilemma then is deciding which is A_S^2 and which is A_R^2 .

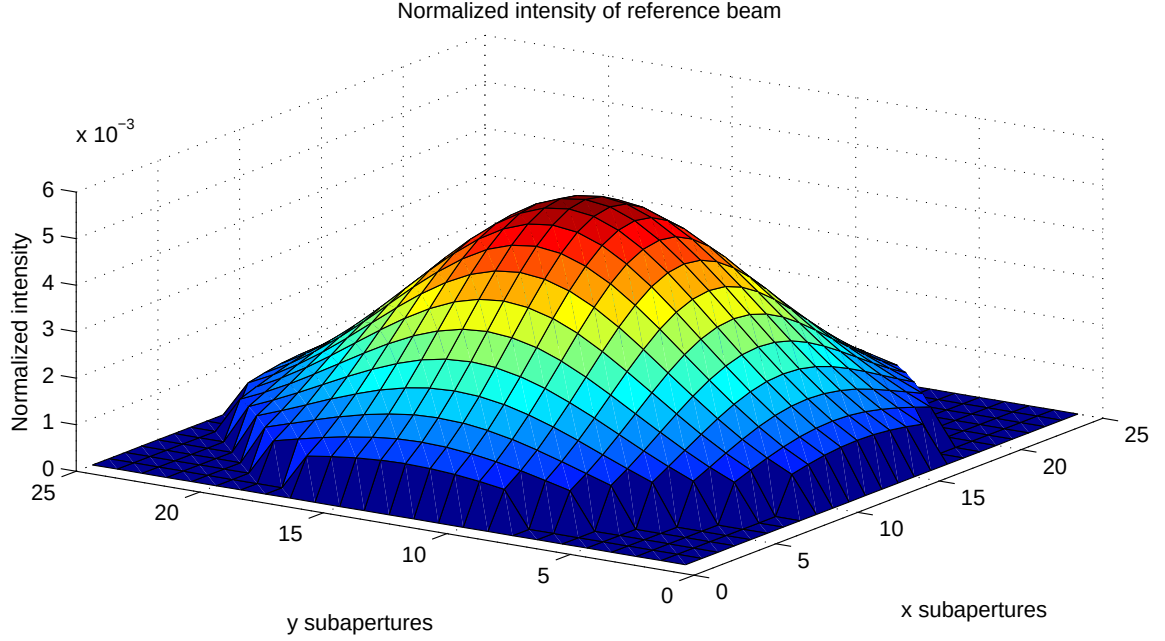


Figure 52: Normalized reference intensity

The first step to accomplish this is to take a calibration of the amplitude of the reference field and normalize this field by the energy of the reference field. Figure 52 shows the idealized Gaussian shape of the reference amplitude for the simulation used throughout this project. The Gaussian shape is caused by the distribution of energy from a finite-sized core of the fiber [44]. This is the shape that the reference field will take so that if the total energy of the reference field is known, the intensity of the reference field at any point can be determined.

With the normalized reference amplitudes in hand, the energy of the entire field is determined as the sum over the entire field of A_S^2 and A_R^2 for a given frame. Conveniently, this is simply the sum over the field of all four interferograms. A portion of this energy is from the reference and a portion is from the signal. Starting from an estimate of having the energy in the reference beam be a factor of the total energy received (say 5 percent to start because the reference beam will be weak until the system locks on) the expected reference amplitudes are generated. Then the energy of each subaperture is determined by summing the interferogram values for that subaperture. If the energy of a subaperture is more than a threshold of twice the energy of the estimate of the reference beam, the larger of the two

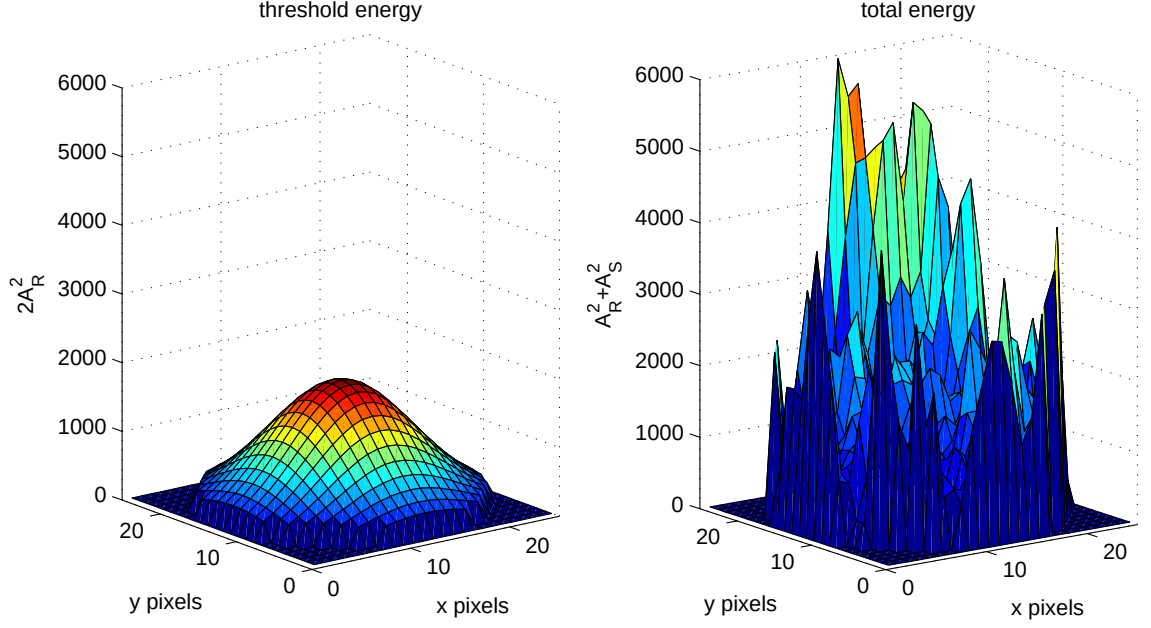


Figure 53: Threshold energy and total energy

values is assumed to be A_S^2 while the smaller is assumed to be A_R^2 .

$$\frac{I_1 + I_2 + I_3 + I_4}{4} > 2A_R^2 \quad (57)$$

$$A_R^2 + A_S^2 > 2A_R^2 \quad (58)$$

$$A_S > A_R \quad (59)$$

If the energy of a subaperture is less than twice the energy of the estimate of the reference beam, the opposite is the case. Figure 53 is an example showing a total amplitude side-by-side with a threshold of twice the reference amplitude. Total amplitudes above the doubled reference amplitude surface have $A_S > A_R$ while total amplitudes below the doubled reference amplitude surface have $A_S < A_R$. For the depicted example in Figure 53, subapertures with $A_S > A_R$ are shown in Figure 54.

Having accomplished this for the entire field, the sum of the reference energy throughout the field is summed and compared with the assumed reference energy. If they are equal (or close) the solution is complete. If they are different, the sum becomes the new reference energy level and a second iteration is accomplished. In simulation, using the factor of total energy determined by the previous frame, the solution resolves to having the refer-

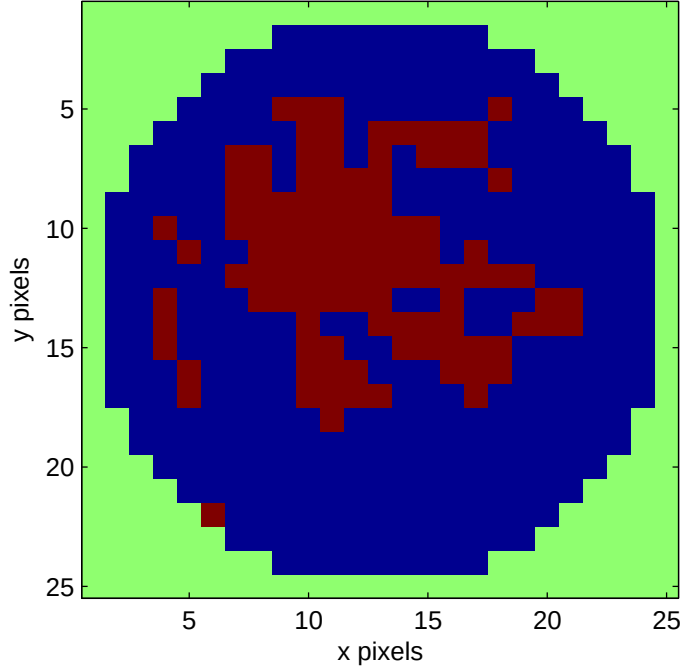


Figure 54: Subaperture segregation: Red = $A_R > A_S$, Blue = $A_R < A_S$

ence energy equate to the estimated reference energy very quickly - usually in one or two iterations and only occasionally three iterations.

Once with a good estimate of the reference energy, A_R^2 is determined from the reference energy and the normalized reference amplitudes. This is believed to be a better estimate of A_R^2 than the solution for A_R^2 determined from the quadratic equation as it utilizes information from the entire field instead of a single pixel, making it less prone to noise. The final portion of the process is to determine A_S^2 from the signal. A_S^2 is determined as the difference between the total energy of a subaperture and the estimated energy of the reference for that subaperture. Total energy for a subaperture is the sum of the interferogram intensities for that subaperture and the reference energy is the previously discussed estimate of the reference energy.

The end result is an estimate of the signal intensity A_S^2 . This is particularly useful in that knowing the intensity of the signal allows the placement of branch cuts into areas of lower intensity.

3.5 Kalman Filter

A Kalman filter accomplishes two goals. The first is that, within the constraints of the fidelity of the system model, a Kalman acts as an optimal *estimator* of the system after a measurement has been provided to the filter. The second is that of a *predictor* of the system between measurements (again, optimally within the constraints of the fidelity of the system model).

While for an AO system sampled at many times the Greenwood frequency, the estimation of the system is generally the more important role, the prediction capabilities of the filter can be important in circumstances such as tracking low-earth satellites where Greenwood frequencies encountered by the system are particularly high due to the speed of the targets across the sky [43].

Restricting ourselves to systems sampled at many times the Greenwood frequency, the implication is that the states of the system are largely unchanged in-between measurements so that the state transition matrix approaches the identity matrix.

3.5.1 Linear Kalman Filter or Extended Kalman Filter. The linearity of system and measurement is a key element in determining whether a linear Kalman filter (LKF) can be utilized in estimating the system or whether the more complicated extended Kalman filter (EKF) must be used.

In the case of a spatially phase-shifted SRI, while the arctangent function which converts interferograms to phases is non-linear, the result of the conversion has been shown to have zero-mean Gaussian properties usable by an LKF. Essentially the spatially phase shifted SRI can be treated as a noisy sensor of ϕ .

For a temporally phase-shifted SRI, however, only one of the four necessary interferograms is produced per sample period. The single interferogram can be used to update an existing state estimate, but the process is non-linear and requires the use of an EKF [41]. The alternative is to save the previous interferograms and treat them as though they all occurred simultaneously (accepting the associated errors with that assumption) even though they did not. Utilizing an EKF has been shown to be advantageous in this circumstance as sample rates lower [41].

3.6 Chapter Conclusions

AO control systems utilizing a fixed-gain proportional integrators already achieve most of the performance benefits of Kalman filter estimation because of their similarity with linear Kalman filters. While augmenting the state variables, varying the proportional integrator gains or even tracking state estimation variances has potential for increasing performance, there is little indication that such improvements would be significant compared to the difficulty in implementing them. The system dynamics and SRI measurement studies necessary to implement any such improvements (other than simply noting the potential for such studies) is beyond the scope of this research and will not be accomplished here.

It is the author's opinion that the computational burden associated with increased system complexities threatens to slow the system down, preventing the performance gains that a more complex system would hope to achieve. In AO (like airplanes) 'speed is good, more is better' and the goal should be to design a system simple enough to work fast while still working well. The conclusion from this chapter is that a standard fixed gain proportional integrator should be used in implementing AO estimation and correction.

IV. When to Unwrap

4.1 Introduction

As discussed in Chapter II, it is important to provide an unwrapped phase to the DM in attempting to correct for atmospheric turbulence. Where in the design to unwrap is an important question to answer when designing an AO system.

4.2 Weak Versus Strong Turbulence AO Systems

4.2.1 Weak-turbulence AO systems. A typical AO system designed to operate in the weak turbulence regime utilizes a Shack-Hartman WFS and a LS reconstructor feeding a proportional integrator as shown in Figure 55.

In such a system, the S-H WFS senses phase gradients $\Delta\phi_x$ and $\Delta\phi_y$ which are restricted only by the physical design of the sensor in range. The S-H WFS works well in weak turbulence conditions because of its simplicity and the fact that there are no significant intensity nulls where the lack of field intensity would cause regions where the WFS gradients had large error variances. The LS reconstructor then generates phases from the phase gradients provided by the S-H WFS. The LS reconstructor generates an unwrapped estimate of the phase of the irrotational portion of the field being seen by the S-H WFS. Since, under weak turbulence assumptions, the field is irrotational, this is a good estimate of the phase of the entire field. The net result is an AO system with a WFS which performs well under weak turbulence conditions using a reconstructor which performs well under weak turbulence conditions.

The phase at this point is the sampled estimate (unwrapped) of the phase of the optical field being seen by the DM (after being corrected by the DM). As such, it is an estimate of the error (or differential) between the phase of the optical field and the phase applied to the DM to correct the optical field. The PI controller integrates a portion of this differential in order to develop the phase $\phi_{DM}(t_i)$ which should be applied to the DM for the next correction given by

$$\phi_{DM}(t_i) = BLS \left(\frac{\partial\phi(t_i)}{\partial x}, \frac{\partial\phi(t_i)}{\partial y} \right) + A\phi_{DM}(t_{i-1}) \quad (60)$$

$$= BU(\hat{\phi}_{WFS}(t_i)) + A\phi_{DM}(t_{i-1}) \quad (61)$$

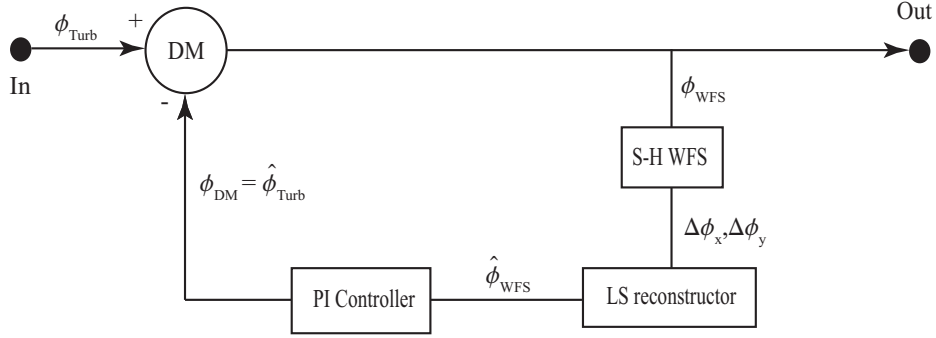


Figure 55: Block diagram of a S-H based AO system

where A and B are the PI gains. $\mathcal{LS}()$ is the least-squares operator, $\frac{\partial\phi(t_i)}{\partial x}$ and $\frac{\partial\phi(t_i)}{\partial y}$ are the x and y phase gradients of the field as measured by the S-H WFS. $\phi_{DM}(t_i)$ is the phase to be sent to the DM and $\phi_{DM}(t_{i-1})$ is the phase that was sent to the DM on the system's previous iteration. $\mathcal{U}()$ is the unwrapping operator and $\hat{\phi}_{WFS}$ is the sampled estimate of ϕ_{WFS} , the phase seen by the WFS.

Once in strong turbulence, however, the situation changes. As previously discussed in Chapter II, S-H WFSs perform poorly in significant scintillation, and LS reconstructors ignore the rotational component of the phase field present in strong turbulence.

4.2.2 Strong-turbulence AO system. As an interferometric-based WFS, the SRI is an excellent choice for systems encountering strong turbulence. However, the phase output of an SRI is fundamentally limited in range to $(-\pi, \pi]$. This restriction is a mathematical artifact of the four-quadrant arctangent function which has a range of $(-\pi, \pi]$. Thus, while an SRI handles scintillation extremely well, includes the rotational phase contributions from branch point effects, and does not need to be reconstructed into phase like the output from a S-H WFS, it still needs to be unwrapped. Unwrapping an optical field is a challenging problem in strong turbulence which is addressed in Chapter V.

Once the output of an SRI is unwrapped, the inclination is to treat the remainder of the system like a conventional design, using Equation (61) and applying the result to the DM in order to correct the wavefront. The problem in utilizing Equation (61) under

strong turbulence conditions is that, unlike the weak turbulence case, being unwrapped does not imply that $\hat{\phi}_{WFS}$ is a sampled estimate of a continuous phase. While being unwrapped implies that wrapping cuts are eliminated, branch cuts connecting the branch points encountered in strong turbulence will still exist. Equation (61) causes these cuts to affect ϕ_{DM} , effectively corrupting ϕ_{DM} .

4.2.3 Uncharted islands. In practice, it has been seen that the iterative integration of the discontinuous $\mathcal{U}(\phi_{DM})$ builds up to the point where not only branch cuts exist in ϕ_{DM} , but wrapping cuts where lines of greater than π difference between adjacent actuator commands extend either from one edge of the aperture to another or islands of unnecessary phase where the cut line extends back to itself as shown in Figure 56. The effect of having wrapped phases on the DM is to have lines of unnecessary 2π phase transitions in the field. These transitions are fundamentally unnecessary and degrade performance because continuous facesheet DMs cannot accurately follow discontinuous phase transitions.

4.2.4 Island persistence. While the wrapping of ϕ_{DM} is an artifact of the discontinuous nature of ϕ_{WFS} , the persistence of wrapping cuts in ϕ_{DM} is a result of the modulo 2π nature of SRI phases. DM actuators and WFS subapertures are matched one-to-one and overlaid on top of each other so that 2π jumps between adjacent subapertures are unsensed and cannot be eliminated. The effect of having wrapping cuts being created and not eliminated is to have the number of wrapping cuts build up over time.

As an example, a simulation was performed where strong turbulence was applied to an AO systems which applied the unwrapper before the PI controller. Log variance was 0.5, $r_0/D_{SA} = 4$, and the sampling rate $f_S = 233f_G$. The performance results for a series of 512 frames are in Figure 57 and show the degradation of the performance for the ‘unwrap then control’ system.

Video of the residual phase clearly shows the buildup of unnecessary phase cuts on the system. While not as demonstrative as the video, snapshots of the residual phase of the ‘Unwrap the control’ AO system at frames 1, 101, 201, 301, 401, and 501 are shown in Figure 58. The first frame is effectively the uncorrected turbulence because the loop has not

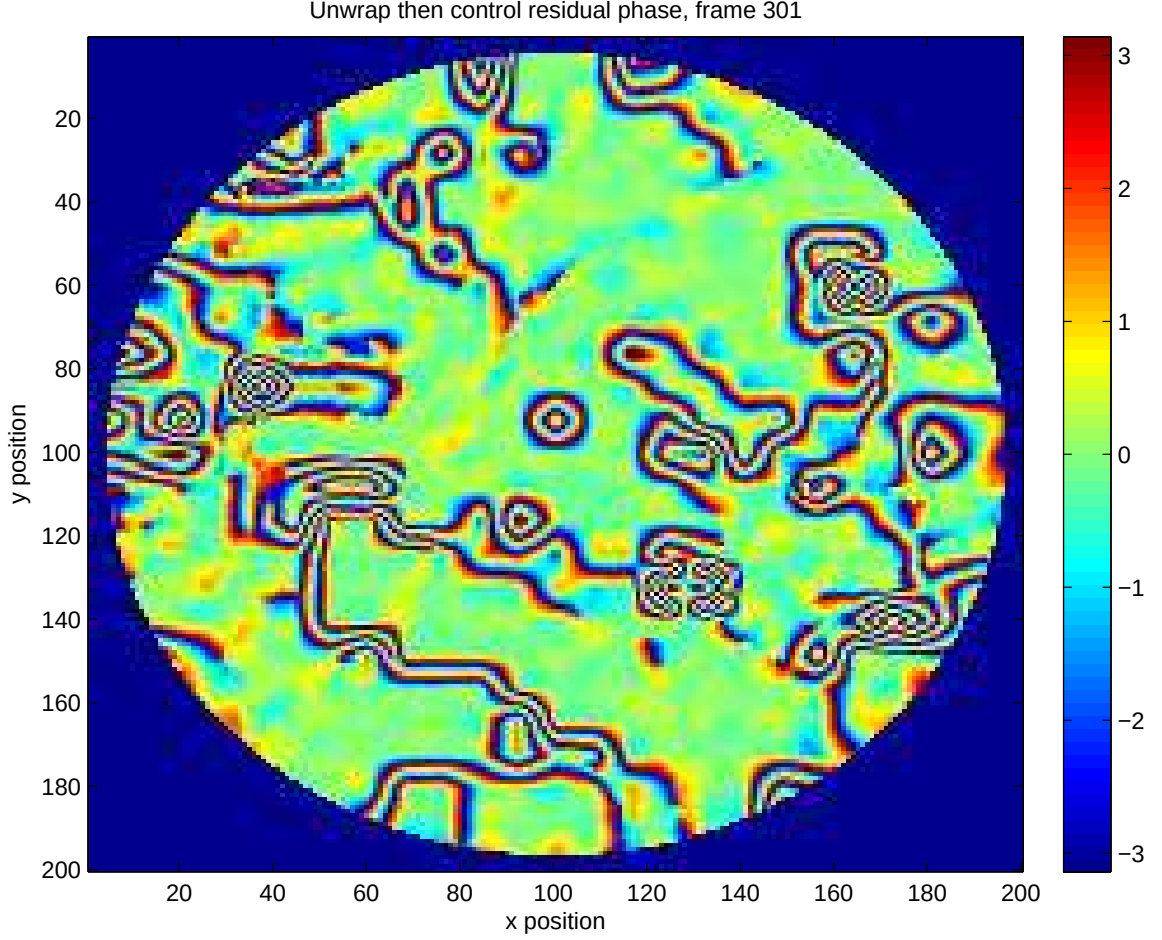


Figure 56: Example residual phase showing 2π islands. Lines of blue/red indicate where the DM is making 2π transitions. Note the isolated circle just slightly above center. The circular shape and the double transition (blue-red-green-blue-red-green) indicate that a single DM actuator is displaced 4π from what it should be.

closed. Subsequent snapshots show the progression of wrapping cuts from less significant to more severe.

4.2.5 Dealing with the problem.

4.2.5.1 Oversampling. The first of several solutions is to oversample the field. That is, having 2^2 or more SRI subapertures for each DM actuator. This allows the SRI to have additional observability on DM induced phase cuts as the 2π differential is spread over several subapertures. This makes it much more likely that the 2π islands can be detected and eliminated. Oversampling the field has been shown to work [36], and

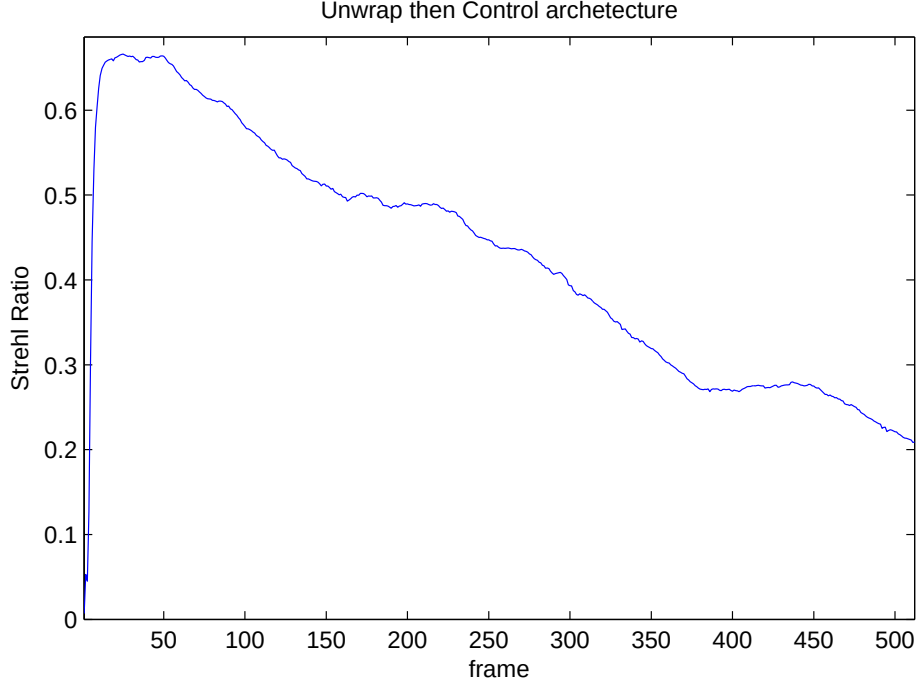


Figure 57: Strehl ratio performance of ‘unwrap then control’ design

should work particularly well when the oversampling is higher, say with 4^2 or 8^2 WFS subapertures per DM actuator. Oversampling reduces the SNR of the WFS, however, and the computational burden associated with doubling (or more) the number of subapertures in the WFS may be a problem and promises to be even more significant as designers create DMs with increasing numbers of actuators, causing the number of WFS subapertures to grow proportionally.

4.2.5.2 Applying a leaky integrator. A second method of avoiding the problem is by putting a ‘leak’ in the proportional integrator which reduces the error in small increments over time. This makes it much more likely that the 2π islands degrade and disappear in the same manner as they appear. Simulations were performed under similar conditions to Figure 58 for PI gain values of ‘ A ’ set to 1.0 (no leak), 0.99 (leak = 0.01), 0.98, 0.95 and 0.9. The residual at frame 501 for each value of A is depicted in Figure 59. Qualitatively speaking it is apparent that the number of unwrapping cuts is decreasing as the leak value increases. The problem is that increasing the leak, while decreasing wrapping cuts, significantly degrades performance of the system in non-wrapping cut areas. This is

apparent by the orange and red areas of the residual in the snapshots of AO systems with higher leak levels. From a more quantitative viewpoint, Figure 60 compares the Strehl ratio performance for the five leak values. From this figure, reducing A to incorporate a ‘leak’ does not help. While reducing A may help alleviate the buildup of 2π islands, the negative effect of the ‘leak’ on performance is significant.

To further investigate this phenomenon, a limited study was performed at the Starfire Optical Range (SOR). The ASALT laboratory at SOR is set up with an AO system whose block diagram is shown in Figure 61. This system is designed to allow users to incorporate different control structures into an AO system with minimal setup and adjustment of the optical bench. Turbulence r_0 and log-variance can be set by positioning phase wheels and the lens sets used with them. The speed that the phase wheels are turned sets the Greenwood frequency of the turbulence and the wheels are reset to a specified initial position so that subsequent runs encounter repeatable conditions. Different types of WFSs are set up in parallel, so that users can choose the WFS they desire without having to rebuild the optical bench. The user can apply MATLAB functions through the use of a Graphical User Interface (GUI) to control outputs to send to the DM based on WFS inputs. For the study, AO control structures were set up identical to the computer simulation. Figure 62 shows a screen shot of the ASALT lab GUI depicting the control setup.

While quantitative data were also recorded, the best indications of the problems of phase cuts building up in a ‘unwrap then control design’ were the qualitative data taken in the form of videos taken showing the interferograms of the surface of the DM.

The video highlighted the problem and the effects of leak values on the buildup of extraneous DM phase cuts on the AO system. Figure 63 shows the interferogram when the DM is flat. Screen Figures 64 through 68 show the interferogram of the DM surface after approximately 100 frames for $A = 1.0, 0.998, 0.99, 0.95$, and 0.9 respectively.

4.2.5.3 A better solution. The previous solutions treat the symptoms of the problem rather than the problem itself in that they attempt to facilitate the elimination of unnecessary phase cuts on the DM after they have developed, rather than preventing the build-up of unnecessary branch points in the first place.

In order to prevent the initial buildup of extraneous branches, a better solution is to place the unwrapper after the PI controller instead of before. The unwrapper can prevent unnecessary phase cuts and keep performance from degrading. It should be noted that the actuator-to-actuator phase differential must be less than π (excluding whatever phase gradient is being corrected by the system's steering mirror.) Actuator-to-actuator phase differentials of greater than π would be modulus restricted to $[-\pi, \pi)$. This deviates from designs used for weak turbulence conditions and is referred to throughout this research as a 'control then unwrap' design. This has been developed in simulation as shown in Figure 69. The key element here is that the proportional integrator develops a wrapped output. Unwrapping the field as the last step before applying the solution to the DM assures that the phases being applied to the DM are unwrapped. As shown in Figure 70, the Strehl ratio performance of such a system is more jagged or noisy than an 'unwrap then control' design, but does not degrade over time as a 'unwrap then control' system will. The performance is an artifact of the unwrapping solutions jumping around somewhat. While this noise is highlighted in simulation it would likely be less noticeable in actual implementation since the DM cannot change instantly as the simulation does.

4.3 Chapter conclusions

The differences between AO under weak or strong turbulence conditions are highlighted by the need to unwrap the field at the appropriate point in the control design. Designs unwrapping phase before applying the PI control law allow unnecessary phase cuts to build up on the DM progressively degrading performance. This problem has been demonstrated in both simulation and laboratory experimentation.

Solutions to the problem of the buildup of extraneous phase cuts include oversampling the field or including a leak factor to help tear down the extraneous phases. Both solutions have problems. Oversampling increases computational complexity, potentially degrading performance if sampling rates are reduced. Applying the PI control law using the wrapped phase estimates from the WFS and then using an unwrapper to unwrap the resulting output prevents the buildup of extraneous phase cuts. This will work under both under weak and strong atmospheric turbulence conditions. The conclusion from this chapter is to use a

‘control then unwrap’ control architecture in designing an AO system for operation under strong turbulence conditions.

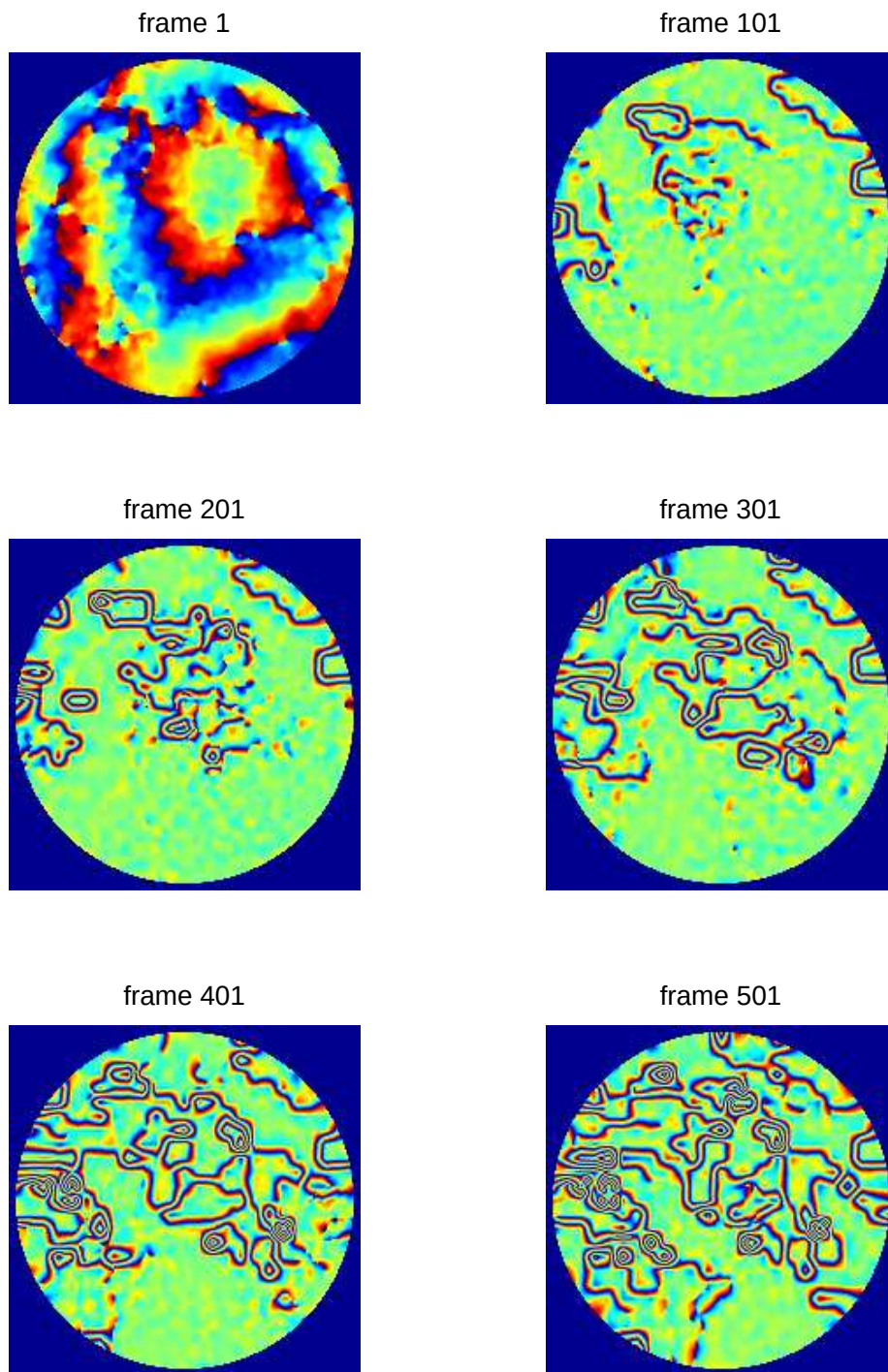


Figure 58: Buildup of phase cuts in ‘unwrap then control’ AO

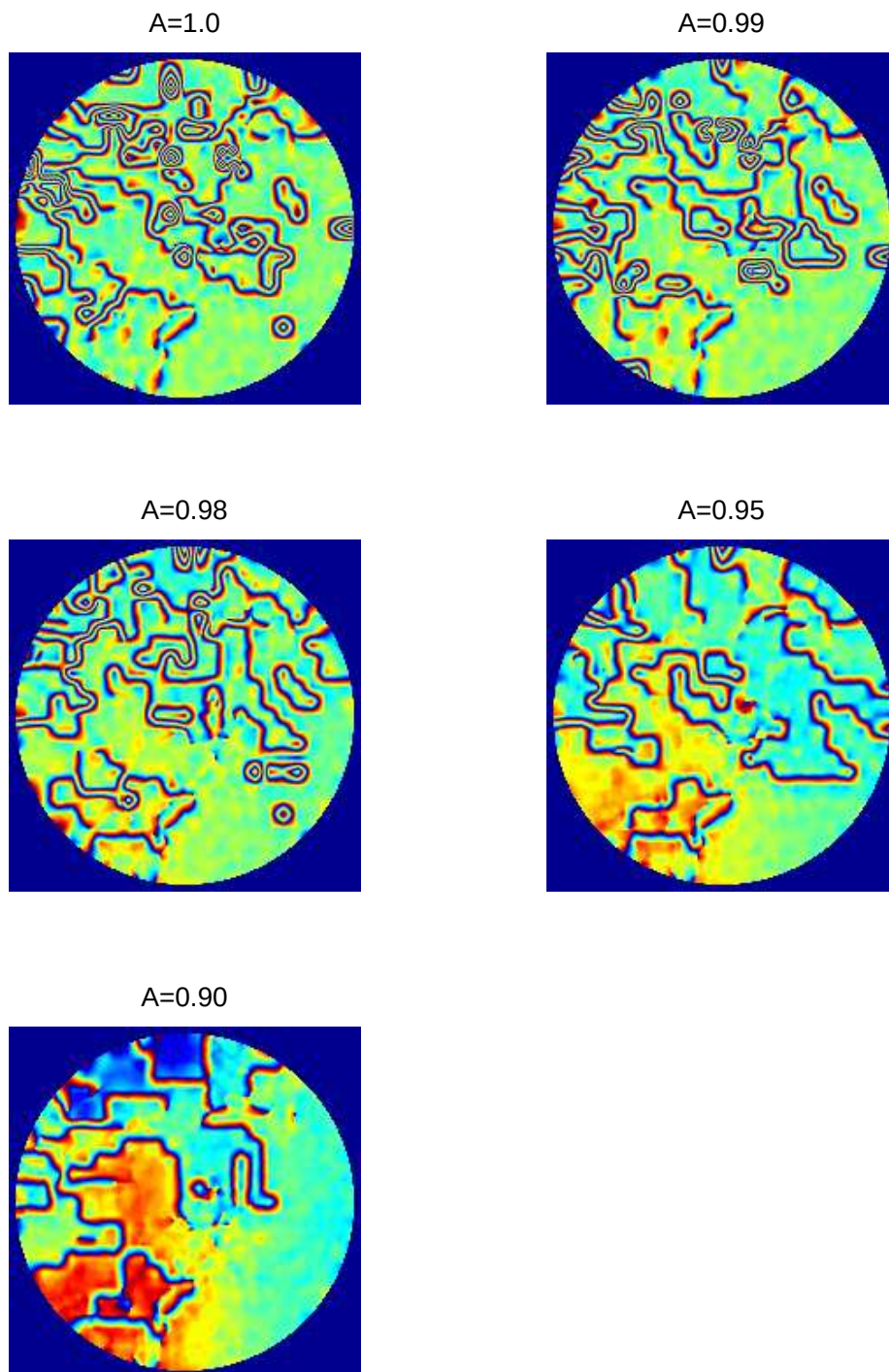


Figure 59: Residuals after 500 frames for varying levels of A

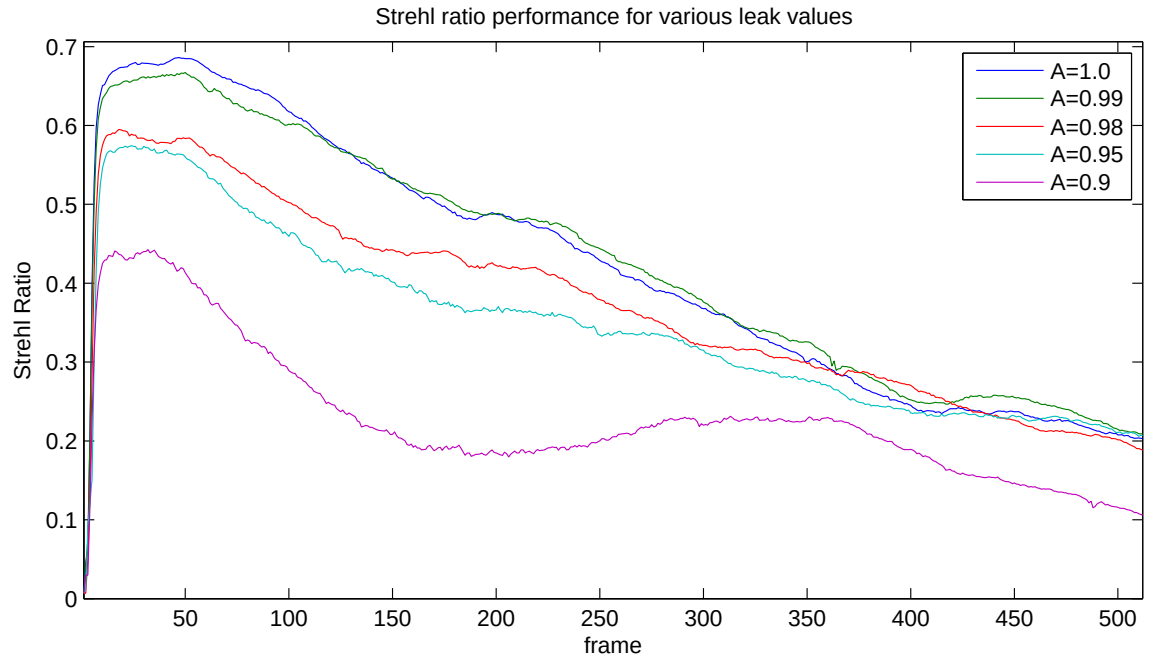


Figure 60: Comparison of various leak levels on system performance

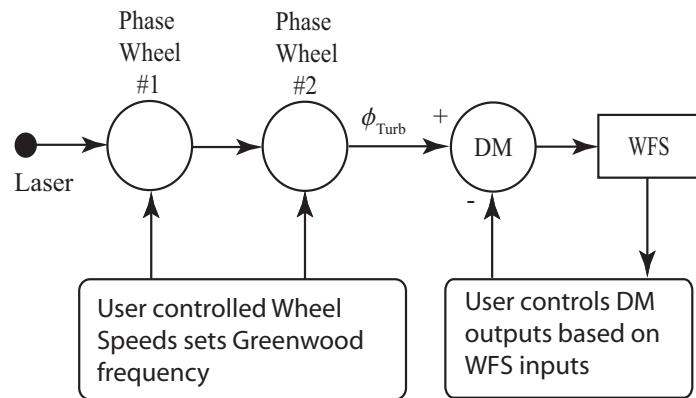


Figure 61: Block diagram of ASALT lab setup

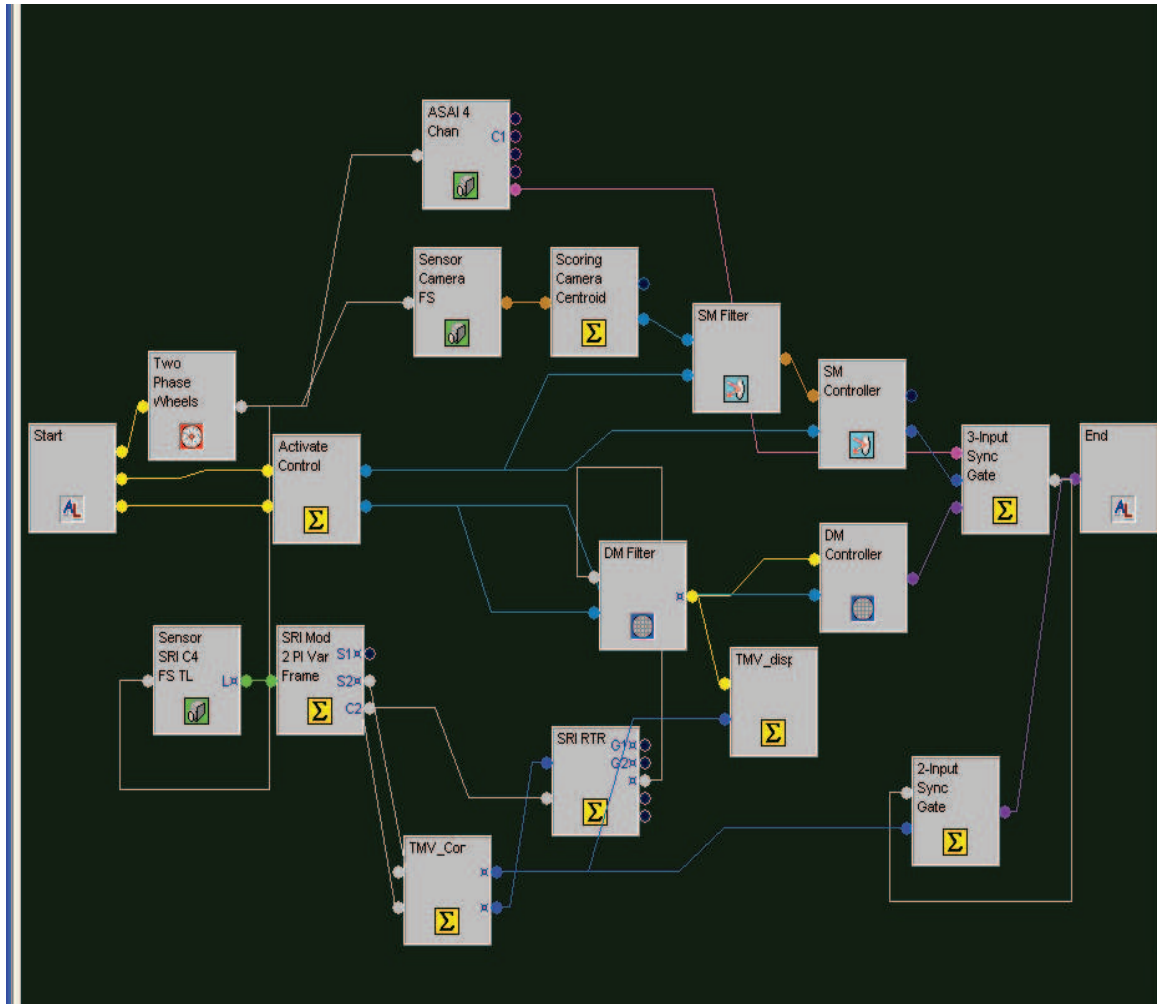


Figure 62: Screen shot of ASALT lab graphical user interface

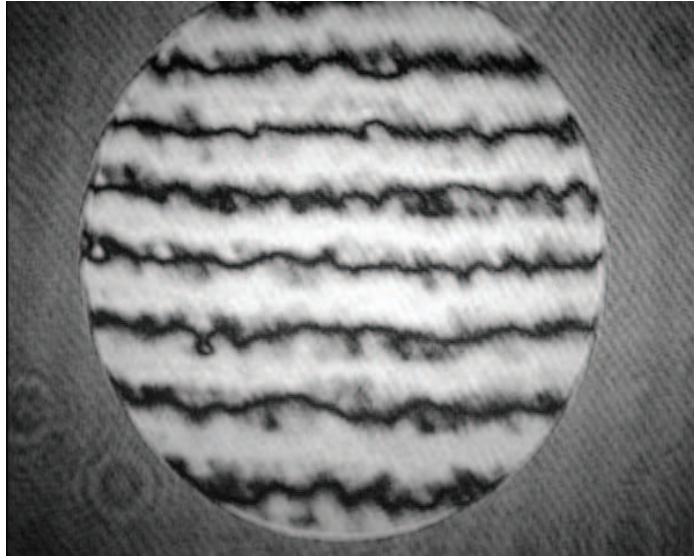


Figure 63: Interferogram showing flat DM (open-loop)

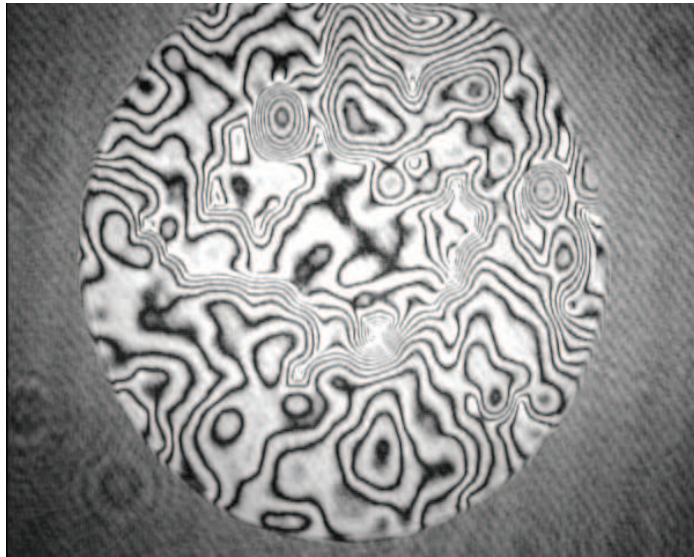


Figure 64: Interferogram of DM after 100 frames in AO system with $A = 1.0$

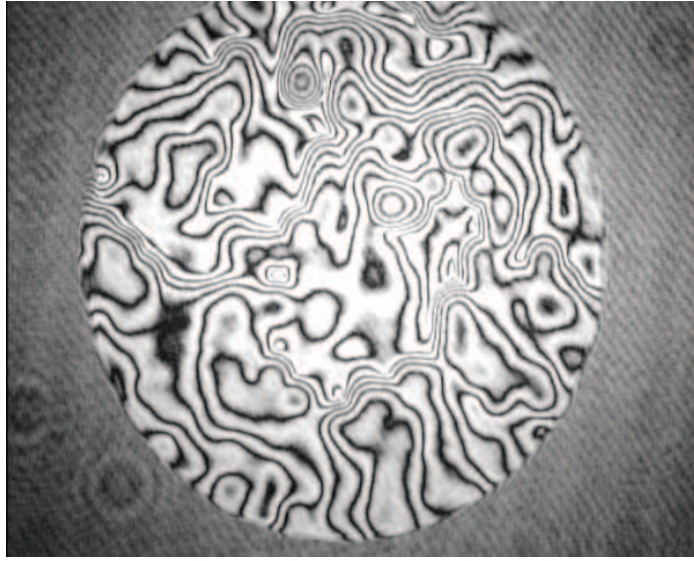


Figure 65: Interferogram of DM after 100 frames in AO system with $A = 0.998$

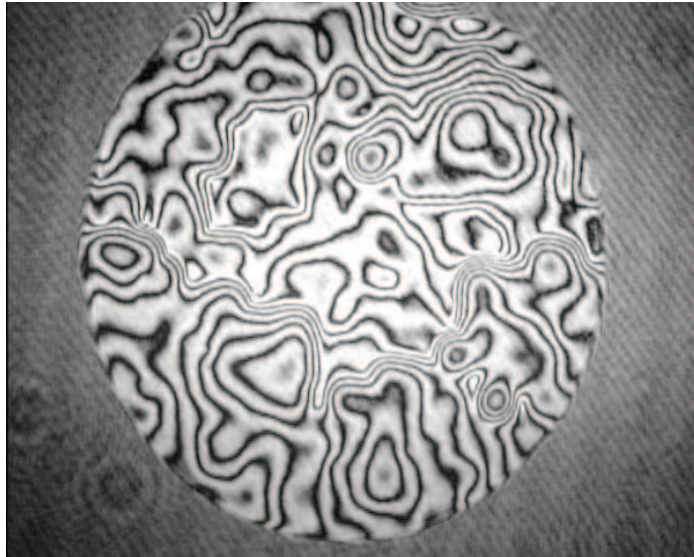


Figure 66: Interferogram of DM after 100 frames in AO system with $A = 0.99$

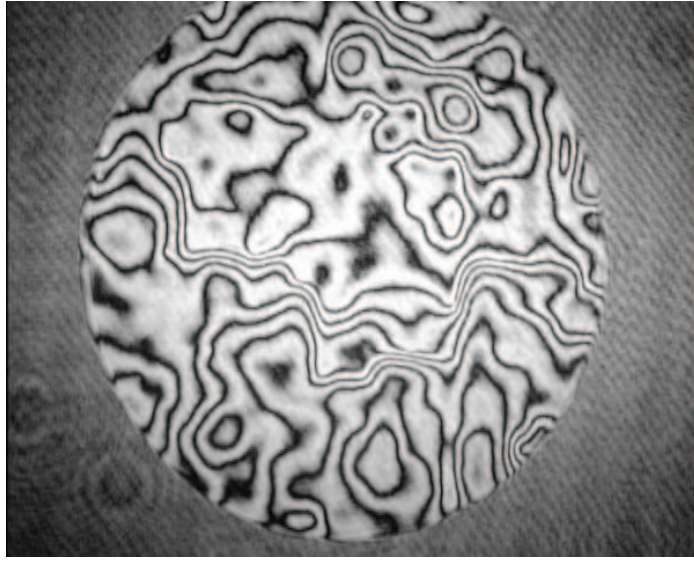


Figure 67: Interferogram of DM after 100 frames in AO system with $A = 0.95$

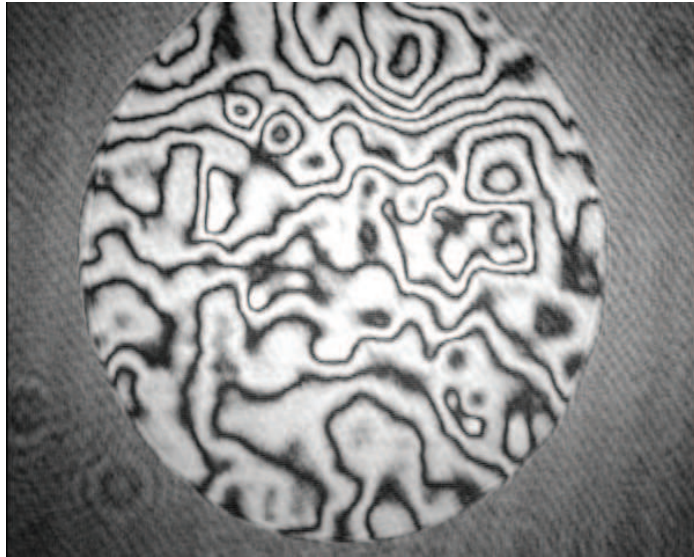


Figure 68: Interferogram of DM after 100 frames in AO system with $A = 0.9$

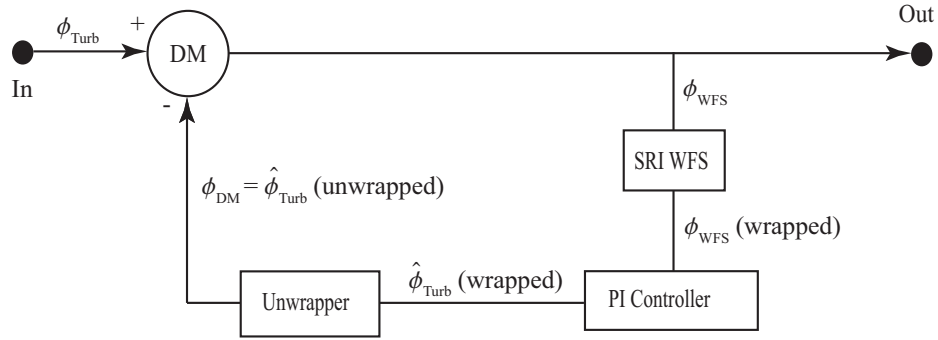


Figure 69: Block diagram of a SRI based ‘Control then Unwrap’ AO system

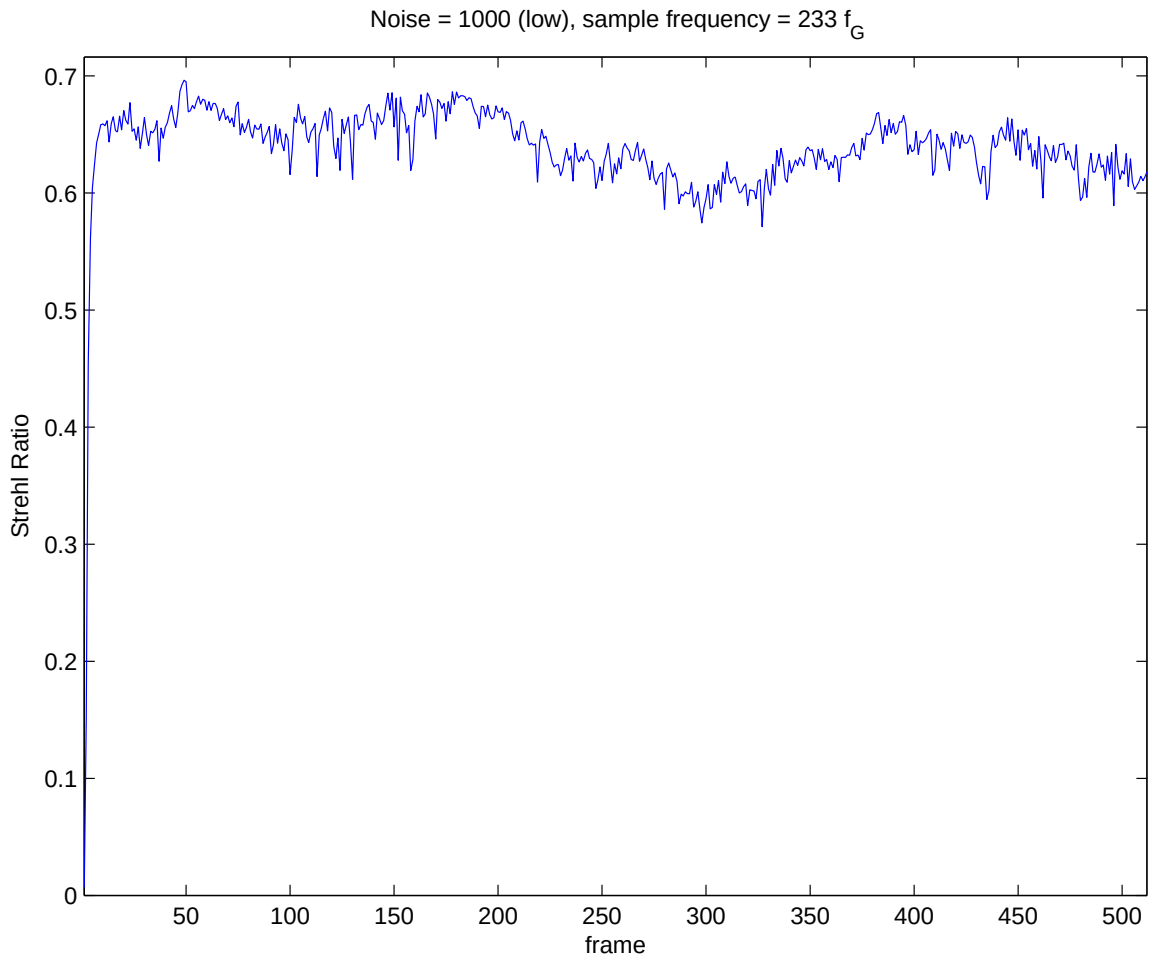


Figure 70: Performance of ‘Control then unwrap’ design

V. Optical phase unwrapping in the presence of branch points

5.1 Introduction

Unwrapping an optical field in the presence of branch points is a significant challenge when designing an AO system to operate in strong turbulence. Under weak turbulence, AO systems utilizing Shack-Hartmann WFSs and least-square reconstructors are unwrapped as part of the reconstruction process. The result is a smooth phase, implementable by a DM. With strong turbulence, branch points require branch cuts which complicate phase unwrapping. These unavoidable 2π lines of discontinuity degrade AO system performance when correcting with a continuous surface deformable mirror due to the inability of the mirror to fit the required discontinuous phase [16]. Branch cut placement, however, affects the amount of degradation and branch cuts can be placed between branch points in many different ways. As previously published by the author [42], this chapter proposes a non-optimal but effective and implementable phase unwrapping method for optical fields containing branch points which places branch cuts where their negative impact on system performance is minimized.

5.1.1 Phase Cuts. Phase cuts, degrade system performance because the DM cannot change shape abruptly and instead changes smoothly between actuators in attempting to match a phase cut. Regions between samples on either side of a cut are poorly corrected by the DM because the DM cannot emulate a cut precisely and will ramp from the commanded level on one side of the cut to the level on the opposite side of the cut. As such, it is advantageous to eliminate phase cuts wherever possible and keep them short and through areas of low illumination when they cannot be eliminated.

For the purposes of this research, a phase cut will be considered as anywhere there is a difference of more than π between adjacent pixels. Throughout this chapter, phase cuts are depicted in figures by lines. ‘x’s and ‘o’s in figures indicate the location of positive and negative branch points, respectively. The line colors are usually white but may vary from figure to figure in an effort to keep them distinct from the background.

5.1.2 Wrapping Cuts. Phase cuts take two forms, wrapping cuts and branch cuts. A wrapping cut is only due to the field being wrapped and proceeds from one edge

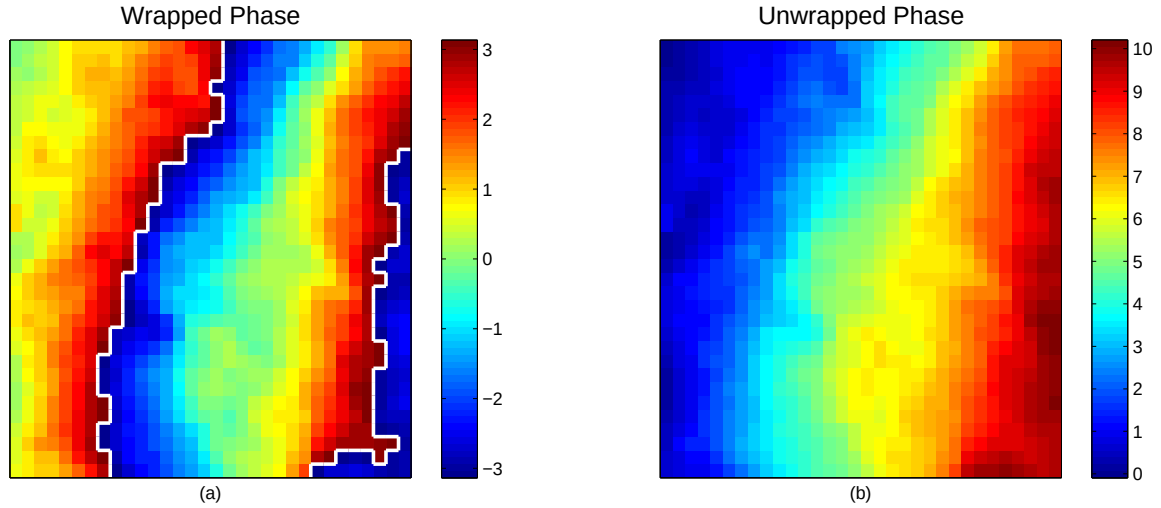


Figure 71: (a) Wrapped phase with only wrapping cuts. (b) Unwrapped version of (a). Note that the unwrapped phase is smooth.

of the optical field to another. It can be eliminated by adding or subtracting an integer multiple of 2π to the field on one side of the wrapping cut. Cuts which form a closed path within the field are also unwrapping cuts and can be eliminated similarly by either adding or subtracting an integer multiple of 2π to the interior or exterior of the cut path. As an example, Figure 71 depicts a wrapped and unwrapped phase. Note that the wrapped phase is limited in range to $[-\pi, \pi)$ while the unwrapped phase is not.

5.1.3 Branch Cuts. Figure 72 shows a phase with both wrapping and branch cuts. Unlike wrapping cuts, branch cuts do not extend across the entire field (or in a closed path) having at least one end terminating at a branch point [17]. They either connect branch points of opposite polarity or connect a branch point with the edge of the optical field (in effect placing a branch point of opposite sign just off the field at that point). By terminating at a branch point, they compensate for the non-zero curl of phase differential around the branch point. In a closed path around a single branch point, the phase differentials integrate to $\pm 2\pi$. As the line integral crosses the branch cut, however, $\mp 2\pi$ is added so that the closed line integral sums to zero as it would if there were not a branch point within the closed path. Branch cuts can be placed in a variety of ways, all of which will still compensate for the non-zero curl of branch points in the phase. Two examples of phase cut placement are shown in Figure 73. The poor unwrap is created by simply unwrapping the field from left

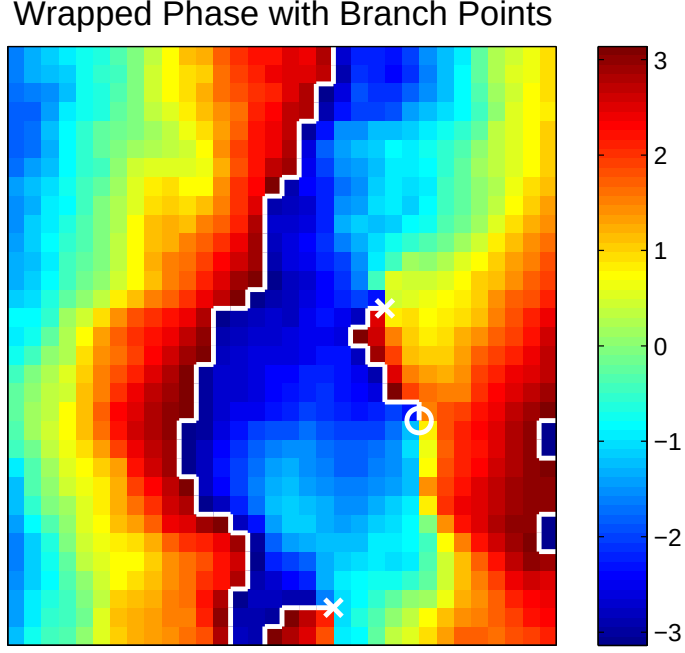


Figure 72: Wrapped phase with both wrapping and branch cuts. If this phase were to be unwrapped, it would not be smooth.

to right. The minimum cut distance unwrap was manually created to minimize the length of the branch cuts.

5.1.4 Least-Squares Unwrappers. Least-squares (LS) unwrappers are very common methods of estimating the unwrapped phase of an optical field in AO systems designed for weak atmospheric turbulence [18]. There are two types, weighted and unweighted.

5.1.4.1 Unweighted LS Unwrappers. For an $N \times N$ array of phases, an unweighted LS unwrapper is developed as

$$\begin{aligned}
 \mathbf{G}\phi &= \mathbf{s} \\
 \mathbf{G}^T\mathbf{G}\phi &= \mathbf{G}^T\mathbf{s} \\
 (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{G}\phi &= (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{s} \\
 \phi_{\text{LS}} &= (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{s},
 \end{aligned}$$

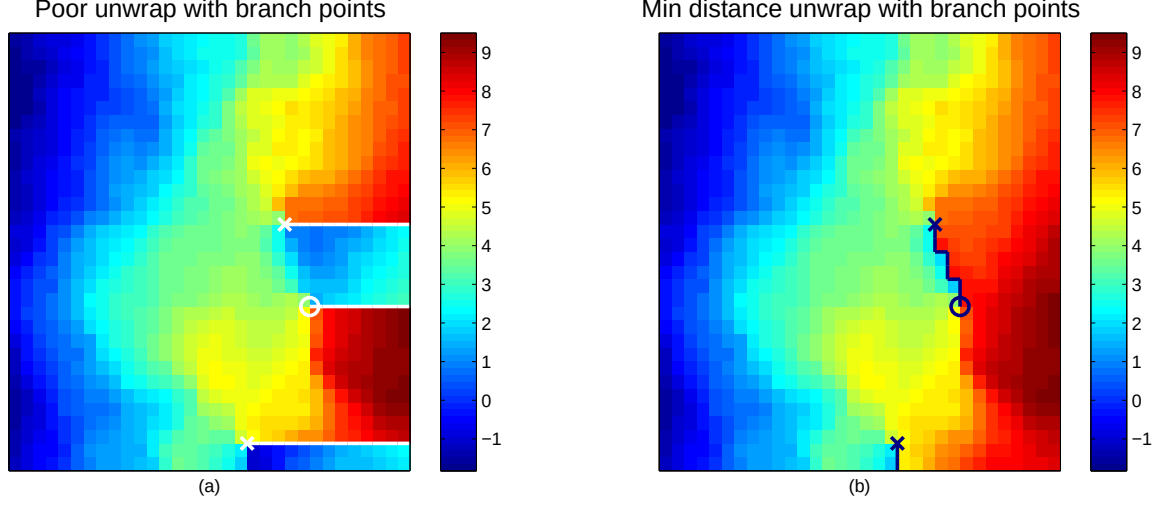


Figure 73: Poor unwrap and minimum cut distance unwrap of phase field with branch points.

where $\mathbf{G}$ is a $2N(N-1) \times N^2$ transformation matrix that converts the N^2 vector of phases ϕ into a $2N(N-1)$ vector of phase differentials in the x and y directions $\mathbf{s}$ and the inverse notation is taken to be the pseudo-inverse. In weak turbulence, $\mathbf{s}$ is most commonly the phase gradients provided by a Shack-Hartmann WFS. If actual phases ϕ_{Tot} are available, the phase differentials $\mathbf{s}$ are developed as $\mathbf{s} = \mathcal{W}(\mathbf{G}\phi_{\text{Tot}})$ where $\mathcal{W}()$ indicates the wrapping operation of limiting the differentials $\mathbf{s}$ to some 2π interval. An important point is that while creating an $N^2 \times N^2$ pseudo inverse is computationally daunting, the problem is alleviated somewhat by $\mathbf{G}$ being sparse and fixed for a given AO system. Much of the work can be pre-computed a single time rather than having to be determined in real time during execution.

5.1.4.2 Weighted LS Unwrappers. Weighted LS unwrappers are sometimes used to minimize noise or emphasize certain parts of a field. In a weighted LS unwrapper,

the slopes are weighted before applying the pseudo-inverse as

$$\begin{aligned}
\mathbf{G}\phi &= \mathbf{s} \\
\mathbf{W}\mathbf{G}\phi &= \mathbf{W}\mathbf{s} \\
(\mathbf{W}\mathbf{G})^T\mathbf{W}\mathbf{G}\phi &= (\mathbf{W}\mathbf{G})^T\mathbf{W}\mathbf{s} \\
\mathbf{G}^T\mathbf{W}^T\mathbf{W}\mathbf{G}\phi &= \mathbf{G}^T\mathbf{W}^T\mathbf{W}\mathbf{s} \\
\mathbf{G}^T\mathbf{W}^2\mathbf{G}\phi &= \mathbf{G}^T\mathbf{W}^2\mathbf{s} \\
(\mathbf{G}^T\mathbf{W}^2\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W}^2\mathbf{G}\phi &= (\mathbf{G}^T\mathbf{W}^2\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W}^2\mathbf{s} \\
\phi_{\text{LS}} &= (\mathbf{G}^T\mathbf{W}^2\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W}^2\mathbf{s},
\end{aligned}$$

where $\mathbf{W}$ is an $2N(N-1) \times 2N(N-1)$ diagonal array of weights. It works essentially the same as an unweighted LS unwrapper, but the pseudo inverse cannot be pre-computed because the weighting matrix is not typically constant. This makes a weighted LS unwrapper difficult to implement in real-time systems.

5.1.4.3 LS Unwrappers and the hidden phase. In estimating the phases from the slopes, there is an implicit assumption that the sum of phase differentials is path independent, or that the field is irrotational. As a result, the phase estimate of an LS unwrapper is irrotational. The LS unwrapper does not reconstruct the rotational portion of the phase, which is why the rotational component of the phase is sometimes referred to as the “hidden phase” [15]. This makes a simple LS unwrapper alone a non-optimal choice when compensating for strong turbulence [32, 39].

5.1.5 Non-LS Component of the Field. The non-LS component of the field is the difference between the original field and the output of a LS unwrapper. If the original field is irrotational, the output of the LS unwrapper will be modulo- 2π -equivalent to the original field, and the non-LS component will be non-existent. If the original field has branch points and is rotational, the effects of those branch points will be isolated in the non-LS component. As such it is sometimes referred to as the rotational component [18]. Strictly speaking, the rotational component containing the non-zero curl effects of the field is not unique [9], so it is referred to here as the non-LS component, uniquely identifying it as the

Field intensity overlaid with branch cuts

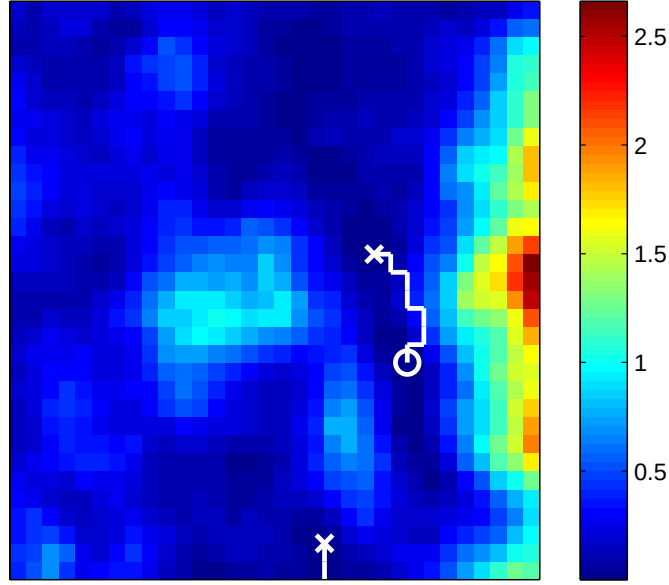


Figure 74: Intensity overlaid by branch cuts using LS unwrapper to eliminate wrapping cuts.

difference between the original field and the output of a LS unwrapper. For the purposes of this research, the non-LS component will be wrapped to a particular 2π range.

5.2 Improved Unwrapper

The first step in unwrapping efficiently in the presence of branch points is generating the LS and non-LS components of the field through the use of an LS unwrapper,

$$\begin{aligned}\phi_{LS} &= LS(\phi_{Tot}) \\ \text{and} \\ \phi_{non-LS} &= \mathcal{W}(\phi_{Tot} - LS(\phi_{Tot}))\end{aligned}$$

where $LS()$ indicates applying an LS unwrapper operation to the vector of wrapped phases ϕ_{Tot} and $\mathcal{W}()$ indicates wrapping the phase to some 2π range.

Wrapping cuts are eliminated by the LS unwrapper, and branch cuts are isolated in ϕ_{non-LS} . Thus total phase ϕ_{Tot} adjusted to remove wrapping cuts while still retaining

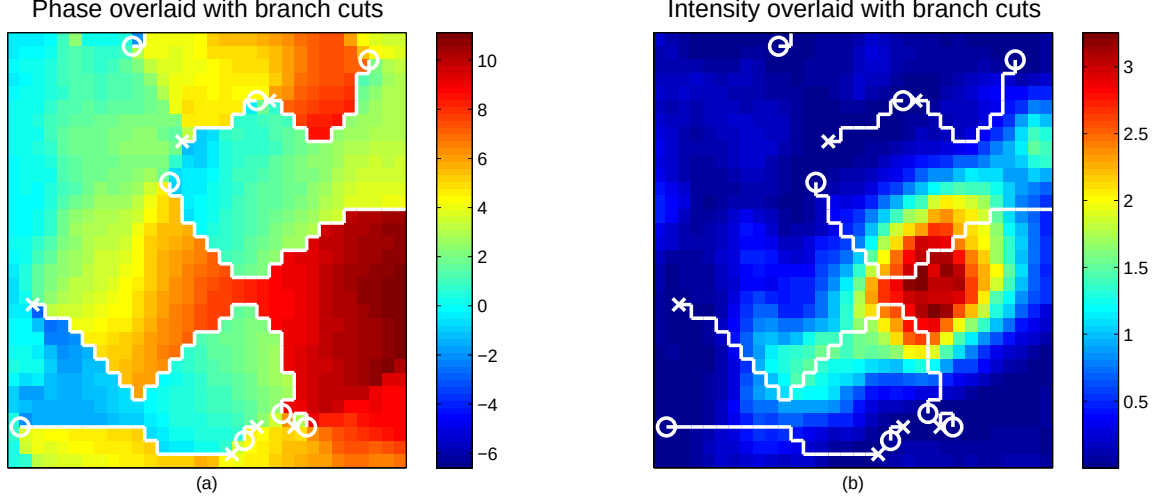


Figure 75: Poor unwrapping, phase and intensity overlaid by branch cuts.

branch cuts can be determined as

$$\mathcal{U}(\phi_{Tot}) = \phi_{LS} + \phi_{non-LS}, \quad (62)$$

where $\mathcal{U}()$ indicates an unwrapping process which removes wrapping cuts (but not branch cuts). While removing any wrapping cuts, this unwrapped result is modulo- 2π -equivalent to ϕ_{Tot} , maintaining both the irrotational and rotational components of the field. This has been covered in several texts [18] and is a common way of including rotational phase effects in the AO systems being developed to operate under strong turbulence conditions [4].

In general, this approach is reasonably effective as shown in Figure 74. Here the field whose phases are in Figure 72 has its wrapping cuts removed by the process depicted in Equation (62). The resultant branch cuts are plotted over the intensity (instead of the phase as in previous figures) to show the effectiveness of the unwrap. In this case the branch cuts are reasonably short and seem to avoid the areas of high intensity, although they may not be optimal.

While generally effective, this unwrap method sometimes gives less appealing results as shown in Figure 75. Here the branch cuts are much longer than they could be and go through areas of high intensity. Admittedly this is the worst realization encountered in the simulation, but poor results are encountered.

Since after unwrapping the LS portion of the phase field ϕ_{LS} is free of phase cuts, the non-LS portion ϕ_{Non-LS} must be examined in order to reduce the impact of phase cuts. Being wrapped, ϕ_{Non-LS} is restricted to some 2π range, say $[0, 2\pi)$. If the range is changed to $[-\pi/2, 3\pi/2)$ then all the points whose phase is in $[3\pi/2, 2\pi)$ would have 2π subtracted from them. The resulting field would be modulo- 2π equivalent to the original field, but would have branch cuts in different positions. The field depicted in Figure 75 is re-depicted in Figure 76 alongside unwraps for the same field with ϕ_{Non-LS} having differing range restrictions. The unwrap with ϕ_{Non-LS} restricted to $[0, 2\pi)$ has terrible branch cut placement. The remaining three realizations depicted are much more reasonable, with the realization created by limiting ϕ_{Non-LS} to $[-\pi, \pi)$ having the lowest normalized cut length. It should be noted that the creation of four realizations is reasonable because the majority of the computational load is in executing the LS unwrapper which only has to be done once.

5.2.1 Unwrapping Metric - Normalized Cut Length. Having developed multiple modulo- 2π -equivalent phase realizations, it is necessary to compare different branch cut placements so that the best one can be chosen. Short cuts through regions of minimal illumination have the least impact on system performance [17]. As such, the metric used in this work is ‘normalized cut length’ which is the line integral of field intensity along any phase cuts divided by the average intensity of the field. It is an indication of what proportion of light in the system is along phase cuts. Since light along branch cuts is erroneously corrected by a continuous facesheet DM, it should be minimized and a shorter normalized cut length is desired.

For a discretely sampled field, normalized cut length is determined by first isolating the phase cuts within the field. This can be accomplished by taking the difference between adjacent pixels first up and down and then side to side. The intensities on either side of the cuts are then summed and divided by two to account for the average intensity along the cuts. Finally, the result is normalized by dividing by the sum of the field’s intensities.

The advantage of normalized cut length is that it can be computed during system execution and is highly correlated to system performance. In order to show the correlation of normalized cut length to system performance, a 256×256 complex ‘Fine’ field is developed

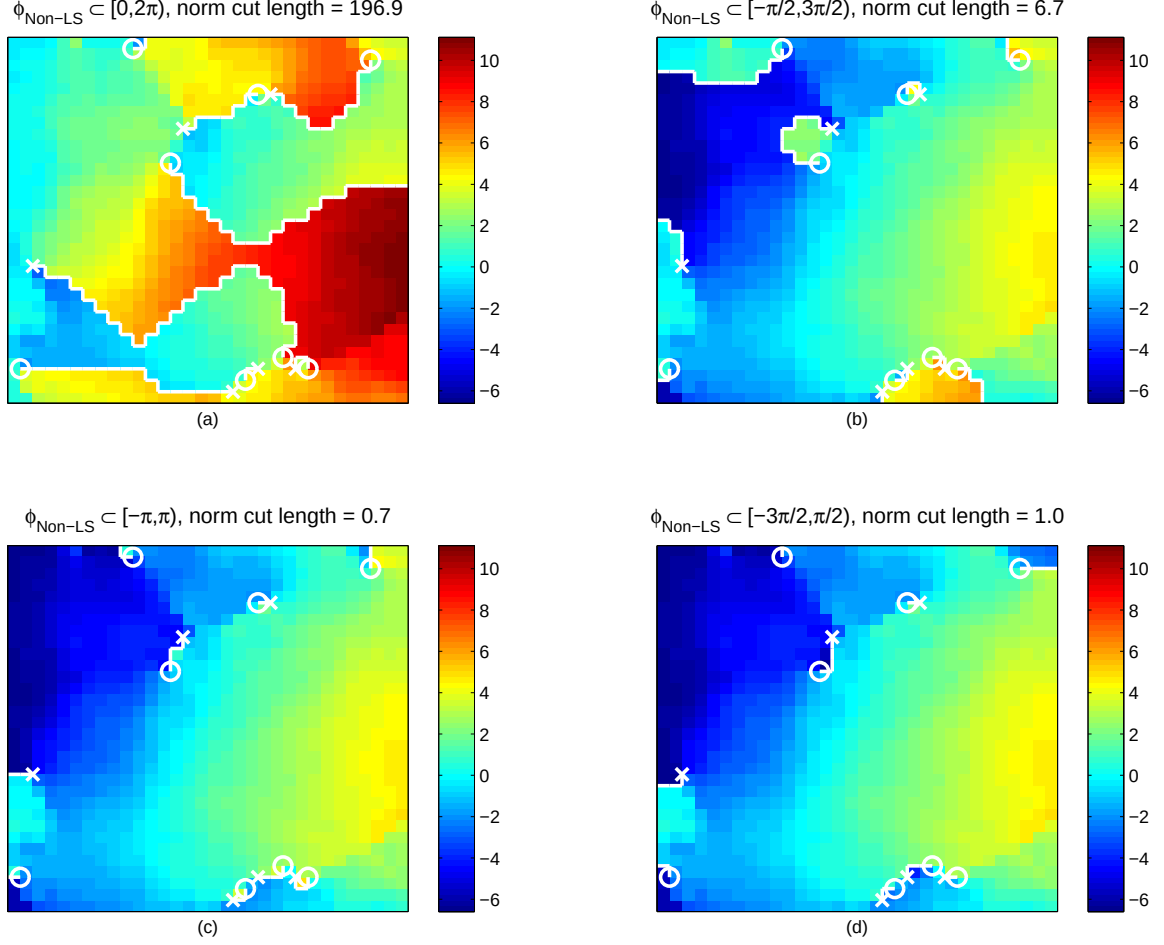


Figure 76: Branch cuts of four different unwrap realizations.

from the 32×32 field by interpolation from the coarser field. Similarly a 32×32 array of phases from an unwrapper is converted into a 256×256 arrays of phases. This models an idealized DM whose surface varies smoothly between actuators. The DM model is translated into the complex domain by $\hat{u} + i\hat{v} = \hat{A} \exp(i\hat{\phi})$ where $\hat{u}$ and $\hat{v}$ are the real and imaginary estimates of the field, $\hat{A}$ is the estimated amplitude of the field and $\hat{\phi}$ is the estimated phase of the field. The field-estimation Strehl ratio can then be computed as

$$S = \frac{\left| \sum_{a_1=1}^N \sum_{b_1=1}^N F_{a_1 b_1} E_{a_1 b_1}^* \right|^2}{\sum_{a_2=1}^N \sum_{b_2=1}^N F_{a_2 b_2} F_{a_2 b_2}^* \sum_{a_3=1}^N \sum_{b_3=1}^N E_{a_3 b_3} E_{a_3 b_3}^*}, \quad (63)$$

where F is the ‘Fine’ field, E is the estimated DM field and $*$ is the conjugation operator [35].

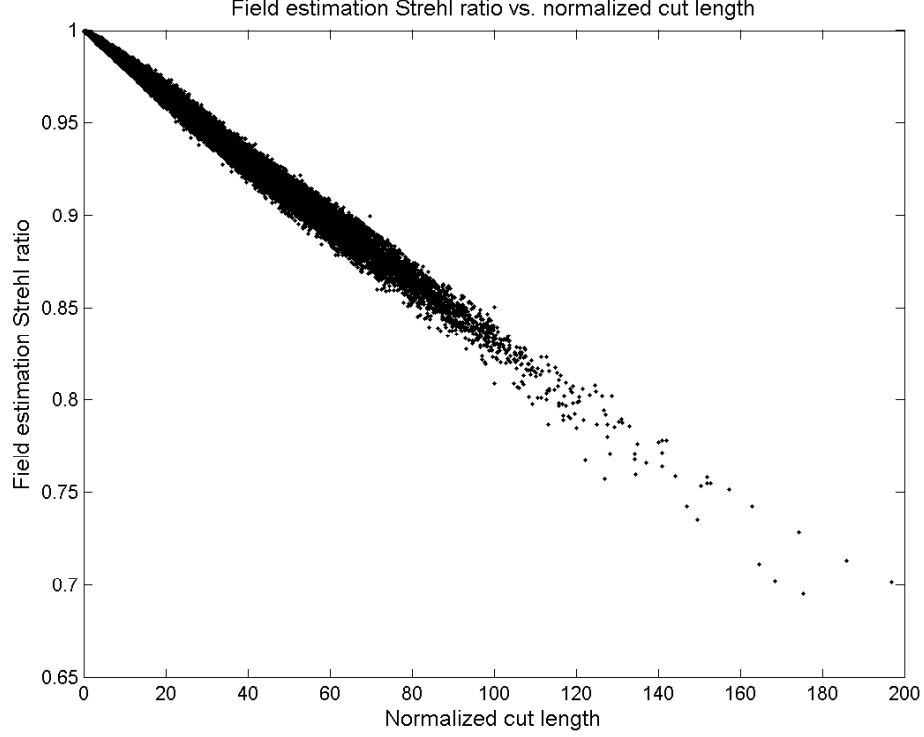


Figure 77: Field estimation Strehl ratio versus integrated cut intensity.

By developing both the ‘Fine’ field and estimated DM field by interpolating from the same coarser field, degradations in the field estimation Strehl are isolated solely to the effect of phase cuts in the field. This allows direct comparison between normalized cut length and the effect of phase cuts on field estimation Strehl.

Normalized cut length is plotted against field estimation Strehl ratio for the various fields and unwrapping methods examined during this work in Figure 77 and has a correlation of -0.9982 . Normalized cut length is shown to be a good measure of the impact of phase cuts on field estimation Strehl ratio, and thus on system performance.

5.3 Simulation and Results

In order to test the unwrapper, a simulation was created that isolates and unwraps 32×32 sections from a 513×513 optical simulation generated test screen. The section size is arbitrary and alternative sizes could be studied. The simulation was run on two test screens depicting fields with intensity log-amplitude variances of 0.4 and 0.8. Each had a

scaling of 16 pixels per the atmospheric coherence diameter r_0 . The log-amplitude variance of intensity is a measure of the scintillation of the field and a reasonable indication of the turbulence strength [1]. Both variances reflect strong turbulence which would create branch points.

For a given field, the 32×32 window is moved throughout the larger test screen to look at all possible realizations of the test screen. While correlated, this gives a wide variety ($482 \times 482 = 232,324$) of different realizations. Some are benign and all four unwrapped versions have effectively placed branch cuts. Others, as shown in Figure 75 have an unwrapped version where a cut passes through or close to a region of high intensity. This is certainly an unwrapping solution to avoid and justifies creating an unwrapper which can choose the best of four unwrap realizations.

For each of the 232,324 possible realizations, the integrated cut intensity metric is recorded for the four ϕ_{non-LS} ranges. The average and maximum score for all realizations is then determined for each of the four ranges. These data show how the unwrapper would perform if the range was fixed to a particular range. The integrated cut intensity is also recorded for the ϕ_{non-LS} range which gives the lowest score. The average and maximum is determined for this best of four ϕ_{non-LS} ranges and compared against the average and maximum scores from the fixed ranges.

The results of the unwrapper using both unweighted and weighted (by field intensity) LS unwrappers to separate out the rotational component are given in Tables 1 and 2 for log-variance values of 0.4 and 0.8 respectively. Compared to limiting the non-LS component to a single range, the variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' mean normalized cut length is reduced in both cases. Perhaps more importantly, the worst realizations are avoided in a variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrapper so that the maximum normalized cut length is dramatically reduced. The weighted variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrapper has the effect of influencing the LS portion of the field towards the areas of higher intensity. The non-LS portion of the field is then influenced towards the areas of lower intensity and branch cuts are forced into darker portions of the field. While a weighted LS unwrapper has the best performance, the computational cost of a weighted unwrapper is significant (see Section 5.4).

Table 1: Normalized cut lengths for 0.4 log-amplitude variance field

	Non-LS component range	Avg norm cut length	Max norm cut length
Unweighted	$[0, 2\pi)$	7.14	196.9
Results	$[-\pi/2, 3\pi/2)$	6.65	159.2
	$[-\pi, \pi)$	6.73	140.0
	$[-3\pi/2, \pi/2)$	7.00	173.0
	best of four realizations	1.13	37.6
Weighted	$[0, 2\pi)$	1.41	88.6
Results	$[-\pi/2, 3\pi/2)$	1.35	90.9
	$[-\pi, \pi)$	1.40	93.3
	$[-3\pi/2, \pi/2)$	1.44	85.7
	best of four realizations	0.62	16.9

Table 2: Normalized cut lengths for 0.8 log-amplitude variance field

	Non-LS component range	Avg norm cut length	Max norm cut length
Unweighted	$[0, 2\pi)$	11.58	150.6
Results	$[-\pi/2, 3\pi/2)$	11.57	161.4
	$[-\pi, \pi)$	11.42	136.5
	$[-3\pi/2, \pi/2)$	11.51	169.6
	best of four realizations	3.0	43.8
Weighted	$[0, 2\pi)$	3.10	94.3
Results	$[-\pi/2, 3\pi/2)$	3.05	108.7
	$[-\pi, \pi)$	2.95	103.7
	$[-3\pi/2, \pi/2)$	2.98	110.6
	best of four realizations	1.43	17.4

5.4 Comparison to Other Unwrappers

In order to evaluate the worth of the variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' method, it was compared to other unwrappers designed to work with branch points. The other unwrappers are the fixed-range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrap, Goldstein's branch cut placement unwrap method [18], Waveprop's xphase [7], and Fried's smoothphase [16].

The fixed-range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrapper is the same as the variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' but only develops a single unwrap realization instead of choosing the best of four realizations. Goldstein's branch cut placement method attempts to determine minimum length branch cuts that connect branch points [18]. Xphase is a MATLAB unwrapping function from the AOTools MATLAB toolbox. It is designed to work with fields containing branch points and attempt to place branch cuts in low intensity regions of the field [7]. It should be noted that 'xphase' required the 32×32 field to be zero-padded to 64×64 in order to work properly because otherwise the field is considered to be periodic [6]. Fried's smoothphase unwrapper separates the field into rotational and irrotational components by first determining the rotational component (after balancing the number of branch points by adding additional branch points along the edge of the field as necessary). Once separated, the irrotational component can be unwrapped and then recombined with the rotational component of the field [16].

The comparison between unwrapping methods is given in Tables 3 and 4 for log-amplitude variances of 0.4 and 0.8, respectively. Execution time is the time needed to execute an unwrapper in MATLAB on a Pentium 4 CPU (3.2GHz) with 2.0 GB of RAM over the 230,000+ frames tested. While execution times may depend on MATLAB implementation, indications from this simulation are that the variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrapper using an unweighted LS gives the best performance at a reasonable computation burden. The variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrapper using a weighted LS improves performance still more, but at an unreasonable computational burden. The AOTools 'xphase' unwrapper gave slightly improved results compared to a variable-range ' $\phi_{LS} + \phi_{non-LS}$ ' using an unweighted LS unwrapper, but at over six times the computational burden.

Table 3: Normalized cut lengths from various unwrappers, 0.4 log-amplitude variance field

Unwrapping Method	Avg norm cut lngth	Max norm cut lngth	Execution time
Unwtd $\phi_{LS} + \text{fxd-rng } \phi_{non-LS}$	7.14	196.9	10 min
Unwtd $\phi_{LS} + \text{var-rng } \phi_{non-LS}$	1.13	37.6	16 min
Wtd $\phi_{LS} + \text{fxd-rng } \phi_{non-LS}$	1.41	88.6	23.6 hrs
Wtd $\phi_{LS} + \text{var-rng } \phi_{non-LS}$	0.62	16.9	23.7 hrs
Goldstein	1.48	71.9	7.5 hrs
AOTools xphase	0.85	37.6	107.8 min
Fried Smoothphase	4.11	175.5	13 min

Table 4: Normalized cut lengths from various unwrappers, 0.8 log-amplitude variance field

Unwrapping Method	Avg norm cut lngth	Max norm cut lngth	Execution time
Unwtd $\phi_{LS} + \text{fxd-rng } \phi_{non-LS}$	11.58	150.6	10 min
Unwtd $\phi_{LS} + \text{var-rng } \phi_{non-LS}$	2.98	43.8	16 min
Wtd $\phi_{LS} + \text{fxd-rng } \phi_{non-LS}$	3.1	94.3	23.6 hrs
Wtd $\phi_{LS} + \text{var-rng } \phi_{non-LS}$	1.43	17.4	23.7 hrs
Goldstein	3.27	84.7	7.6 hrs
AOTools xphase	1.83	38.0	110 min
Fried Smoothphase	9.33	182.7	19 min

5.5 Impact on System Performance

The purpose of developing an improved unwrapper is to improve the performance of a closed-loop AO system encountering strong turbulence. As such, a 1000 frame closed-loop AO simulation was performed under 0.5 log-amplitude variance strong turbulence in order to compare the effect of the unwrapping on system performance.

With the exception of the log-amplitude variance, simulation conditions were purposely benign in order to isolate the unwrapping as the dominate factor on system performance. The remaining simulation conditions were $r_0 = 4D_{SA}$ where D_{SA} is the diameter of a sub-aperture, sample rate = $223 f_G$ where f_G is the Greenwood frequency of the atmosphere, and average $SNR \simeq 200$. The simulation used a leak-free integrator controlled feedback with a error signal gain of 0.4. The control law was applied immediately before the unwrapper, whose output then went to the DM which is the design advocated in Chapter IV.

System performance using fixed ϕ_{non-LS} range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrappers was compared against using a variable ϕ_{non-LS} range ' $\phi_{LS} + \phi_{non-LS}$ ' unwrapper. Strehl ratio

Table 5: Average Strehl results for 1000 frame simulation

ϕ_{non-LS} range	mean Strehl ratio	‘Best of four’	‘Best of four’
		avg improvement	max improvement
$[0, 2\pi)$	0.5956	7.6%	29.8%
$[-\pi/2, 3\pi/2)$	0.6104	5.0%	41.6%
$[-\pi, \pi)$	0.6198	3.4%	23.0%
$[-3\pi/2, \pi/2)$	0.6024	6.3%	33.4%
best of four	0.6406	N/A	N/A

performance of the various simulations are plotted in Figures 78 and shows how different fixed ranges have different periods of reduced performance. Average results are tabulated in Table 5 as well as average and maximum improvements when using a ‘best of four’ unwrapper. The new unwrapper improved the average Strehl ratio performance between 3.3% and 7.6% against the four fixed ϕ_{non-LS} range unwrappers with considerably less variability. The maximum improvement of the new unwrapper against the four fixed ϕ_{non-LS} range unwrappers was more dramatic, ranging from 23.0% to 41.6%.

As the performance of a fixed-range ‘ $\phi_{LS} + \phi_{non-LS}$ ’ unwrapper is inconsistent, the average improvement of the variable-range ‘ $\phi_{LS} + \phi_{non-LS}$ ’ unwrapper over a fixed-range ‘ $\phi_{LS} + \phi_{non-LS}$ ’ unwrapper is difficult to determine. In order to develop an average improvement, the simulations were extended to 10,000 frames to provide each fixed-range of the ‘ $\phi_{LS} + \phi_{non-LS}$ ’ unwrappers with areas of both good and bad performance. The results of the simulation were put into histograms and then summed to form cumulative distribution functions (CDFs) shown in Figure 79. The CDFs show how the variable range ϕ_{non-LS} unwrapper improves performance. The CDF of the variable range ϕ_{non-LS} unwrapper is shifted to the right when compared to the CDF of the fixed range ϕ_{non-LS} unwrapper. Not only does this indicate improved average performance, but indicates more significant improvement for systems such as laser communication where performance thresholds which inhibit operation below a certain Strehl ratio.

5.6 Chapter Conclusion

In the presence of branch points, unwrapping the phase is a difficult problem. Isolating the rotational component by using a LS unwrapper to separate the field into its LS and non-LS components seems an excellent approach. The wrapping phase cuts of the irrotational

component are automatically eliminated by the LS unwrapper. Altering the range of the rotational component is a simple and effective way of varying the placement of the branch cuts associated with the rotational phase component, and computing the normalized cut length is an effective way of comparing the effectiveness of branch cut placements. Choosing the best of four branch cut realizations not only improves average cut placement but, perhaps more significantly, eliminates the worst cut placements which would significantly degrade AO system performance. The improved unwrapping eliminates regions of degraded performance where previous unwrappers yielded poor branch cut placements. The reduced areas of poor performance not only improves average performance, but may significantly improve systems such as laser communications where falling below a performance threshold causes signal fading.

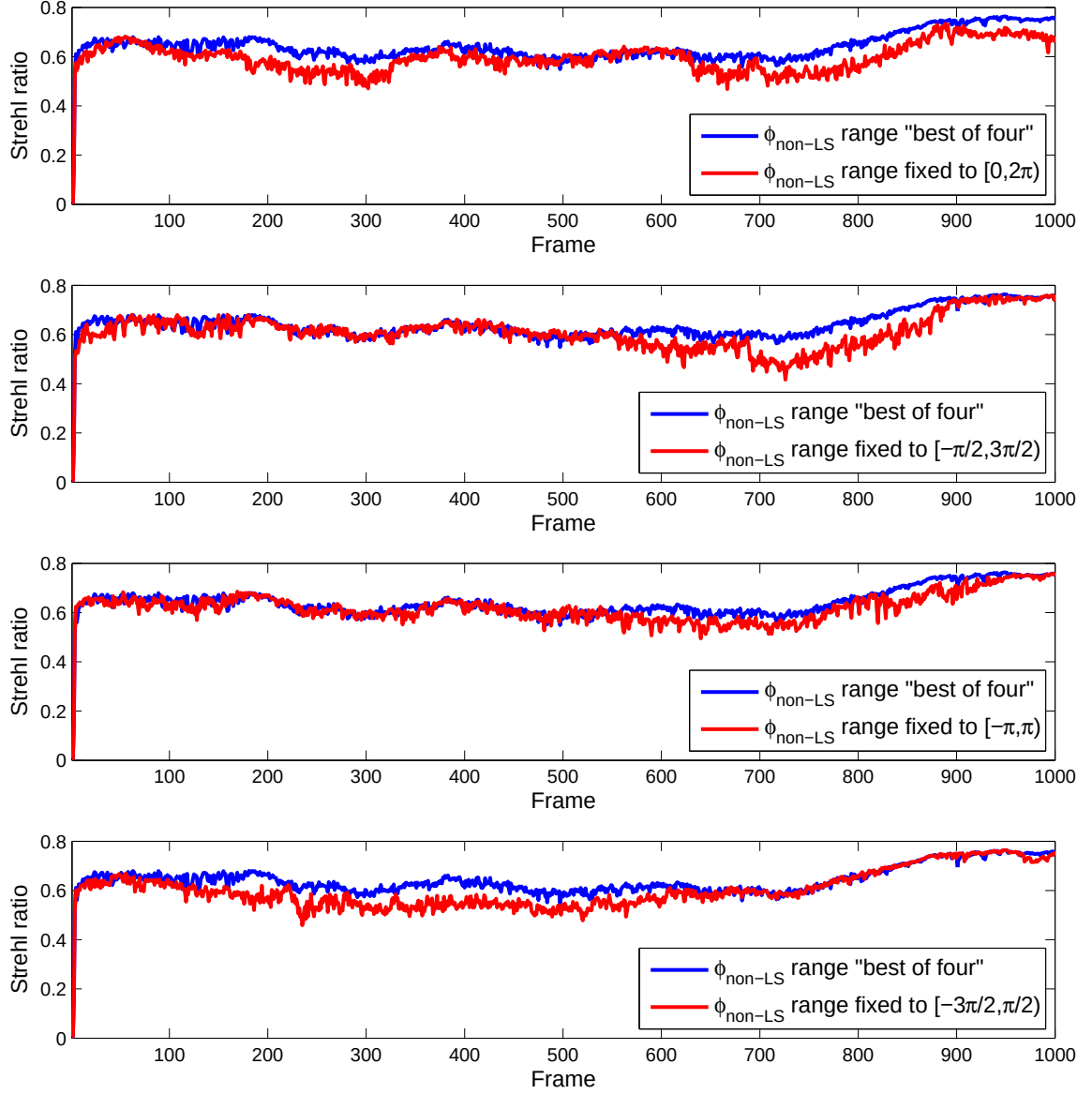


Figure 78: Comparison of closed-loop AO performance between variable and fixed $\phi_{\text{non-LS}}$ range ' $\phi_{\text{LS}} + \phi_{\text{non-LS}}$ ' unwrappers.

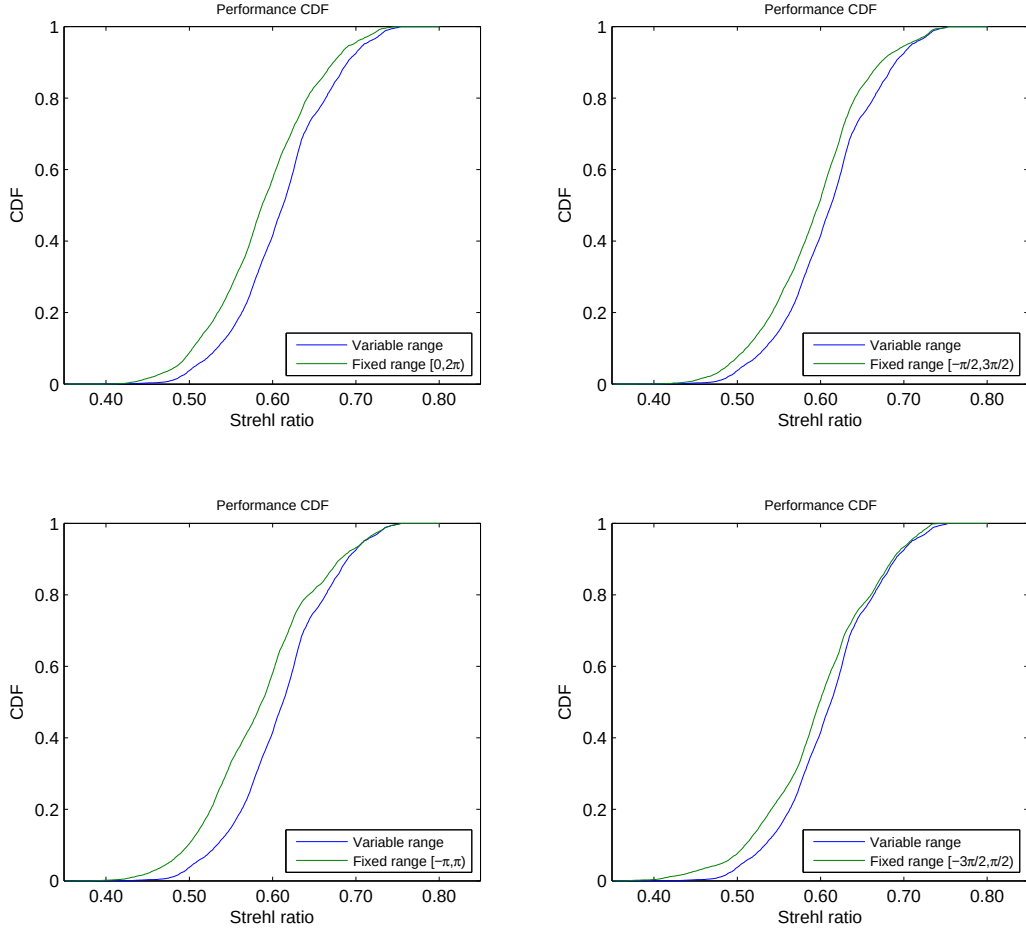


Figure 79: CDF comparisons between variable range ϕ_{non-LS} and fixed range ϕ_{non-LS} unwrappers.

VI. Results

6.1 Introduction

This chapter combines the results of this research into an basic AO control structure and tests this structure in simulation. With the exception of log-variance, the baseline conditions are purposely benign, to allow the parameters to vary one at a time into less benign regions. The amplitude log-variance is 0.5 in all cases to generate the strong turbulence conditions where branch points will be present in the optical field. The simulation is tested against varying r_0 , sample rates, and read noise.

The baseline conditions for the system were $D_{SA}/r_0 = 0.25$, sampling rate $233f_G$, log-variance 0.5 and read noise 1000 (this is explained in section 6.5, but is low). The baseline Kalman gain is 0.4 and is a scalar because there is a single state variable ϕ . The only variation of the Kalman gain is to illustrate the effect of altering the Kalman gain under low-noise, low sample rate conditions.

6.2 Basic AO structure

The basic structure of the AO system under test as developed by this research is shown in Figure 80. The system has WFS subapertures and DM actuators overlaid in a one-to-one mapping as shown in Figure 81.

6.3 Varying r_0

When considering r_0 , the relevant values are the ratio of r_0 to the subaperture size of the system. Figure 82 shows a performance curve over 512 frames for the baseline conditions where $D_{SA}/r_0 = 0.25$. The D_{SA}/r_0 ratio was then varied from 0.1 to 1.0 to establish the

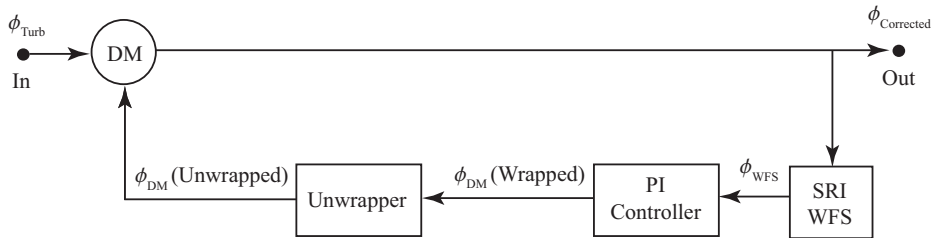


Figure 80: Block diagram of the AO system.

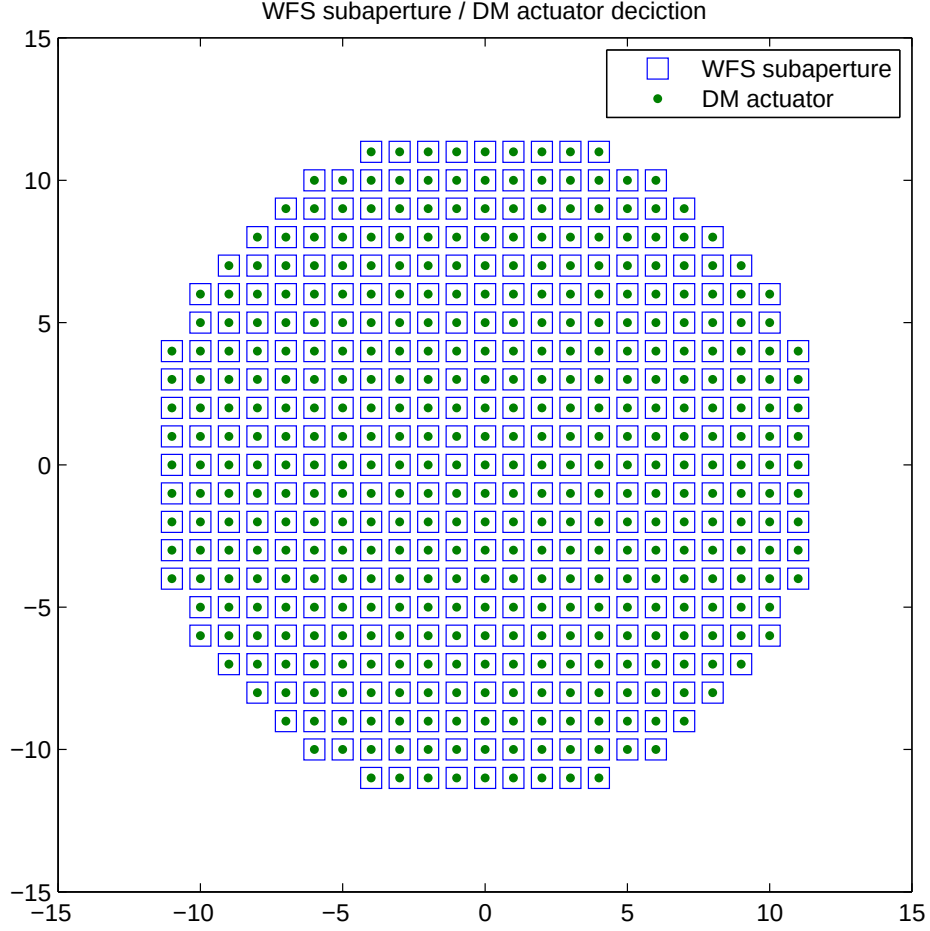


Figure 81: Depiction of WFS subaperture and DM actuator positions.

impact of D_{SA}/r_0 on system performance. The results are shown in Figure 83 where the performance of a system is the average Strehl ratio of the system. As the graph shows, as subaperture sizes decrease, performance improves. Larger subapertures are unable to accurately correct the phase of the wavefront so that as D_{SA}/r_0 increases, performance decreases.

The performance is calculated as the average Strehl ratio of the system for ten different simulation realizations where a realization is defined as a 512 frame simulation generated from a particular random seed. The performance of the system for a particular realization was determined as the average Strehl ratio of the system for frames 101 to 512. Omitting the first hundred frames was intended to allow the system to be completely locked on the

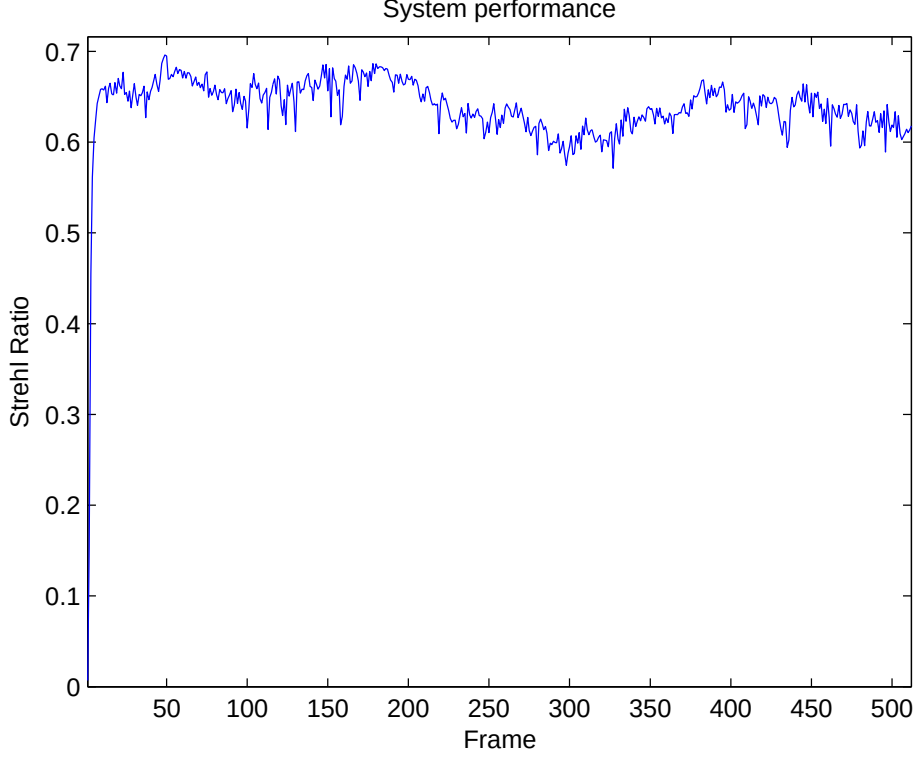


Figure 82: System performance (Strehl ratio) for $D_{SA}/r_0 = 0.25$

turbulence before beginning to determine its performance. This method of determining performance is similarly applied to the other parameter variations in this chapter.

6.4 Varying sample rates

The sample rate of the system, as compared to the Greenwood frequency of the turbulence is an important consideration when designing an AO system. Generally it is best to sample as fast as possible, but sampling fast has implications to signal quality as well as computational burden. The quality of the signal is degraded at the faster sampling rates because the sensor integrates the light over the WFS subapertures for shorter amounts of time. In addition, the sensor is required to read out the data and function at the higher speeds. The computation time is important because it is a prime factor in the delay between a WFS measurement and the application of a modified correction based on that measurement to the DM. The simulation takes the simplistic approach that the computations are

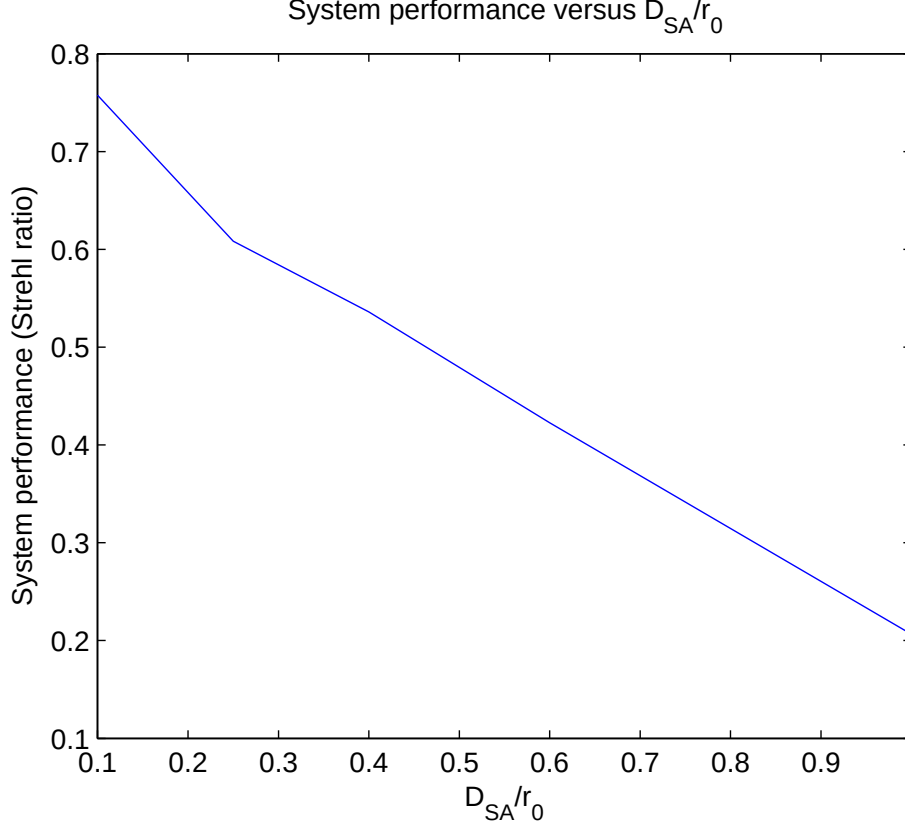


Figure 83: System performance (Strehl Ratio) vs. D_{SA}/r_0

accomplished in one frame which was deemed adequate for this research. The results are plotted in Figure 84 and show the importance of sampling much faster than f_G .

At higher sampling rates, varying the Kalman gain K had little effect on system performance. This is because the system changed very little between samples and the signal from the WFS was fairly clean. Performance in this case was most affected by r_0 and amplitude log-variance. For the lower sampling rates, however, system performance is improved by increasing the Kalman gain K . This is because the system is changing more between samples and the PI controller should take a higher proportion of the WFS input. It is quite likely that increasing the states of the system to include one or more temporal differentials of phase would be effective in these lower sampling rate conditions, but that was not investigated in this research. The effect of varying K at the fixed sampling frequency of 10 times f_G is portrayed in Figure 85. While this is admittedly under low-noise conditions, it highlights the impact of the Kalman gain K at lower sampling rates. Moreover, the

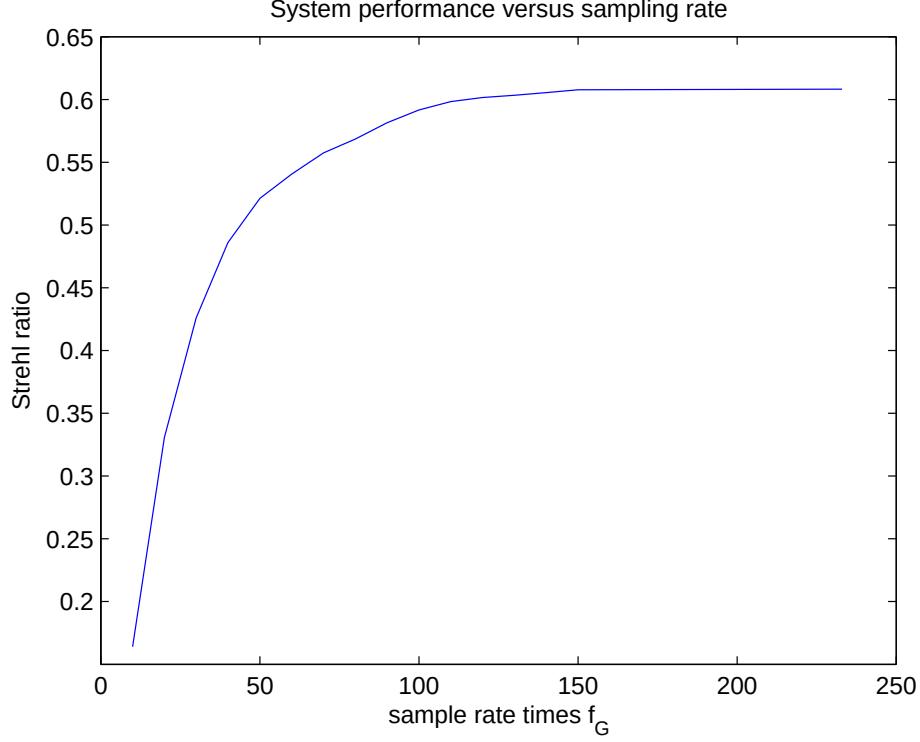


Figure 84: System performance (Strehl Ratio) vs. normalized sampling rate

fact that the performance is increasing for $K \geq 1$ indicates that using additional differential phase states could further improve performance under low-noise, low sample rate conditions. This is covered in more detail in Section 7.3 which covers future work.

6.5 Varying read noise

The final variation is that of the noise of the sensor. Sensor noise is considered as read noise from the sensor and shot noise from the signal. Read noise from the sensor is modeled as a fixed variance Gaussian. Shot noise is well modeled as a Poisson distribution of the number of photons received. For simplicity, the signal is considered to be strong enough (have enough photons) that the Poisson distribution shot noise associated with the signal can be considered Gaussian with a variance equivalent to the intensity (measured in photons) encountered by the sensor. The net result is a single Gaussian noise of strength described by

$$\sigma_{Noise}^2 = \sigma_{Read}^2 + I_{measurement} \quad (64)$$

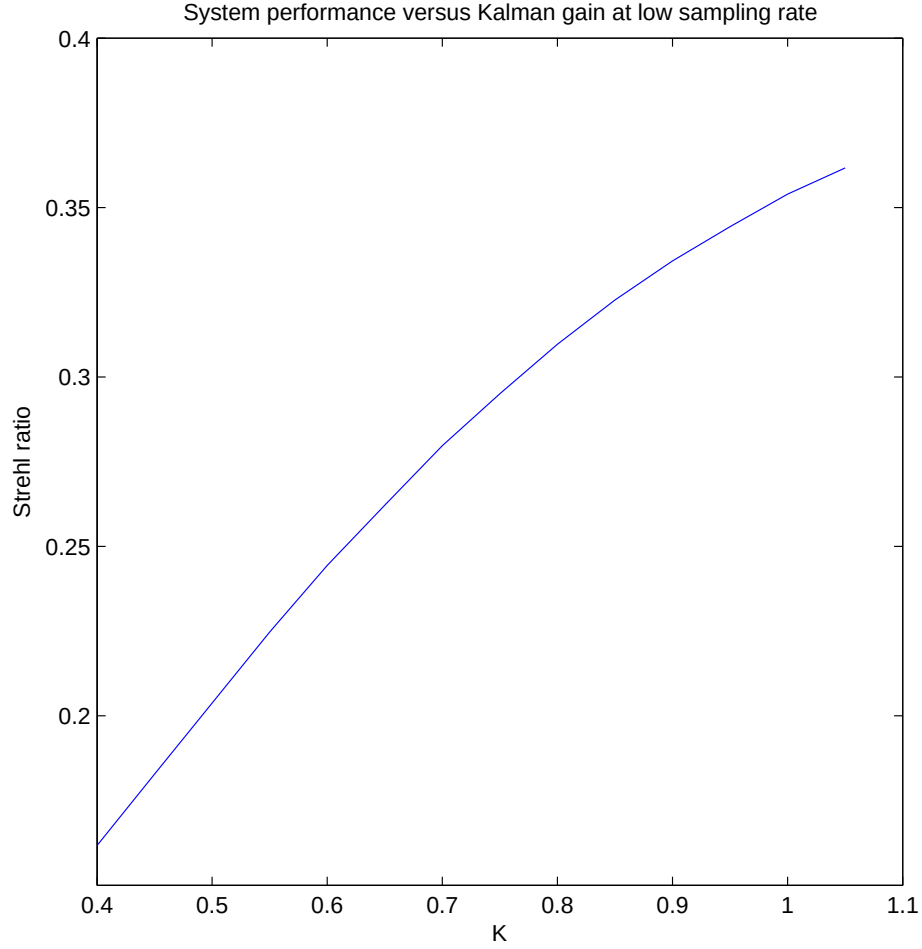


Figure 85: Effect of varying the Kalman gain K for low noise, low sample rate

which is added to the signal.

The field strength of the simulation is such that the read noise is dominant with shot noise contributing significant portions of the noise at only the highest regions of signal intensity. This is shown by Figure 86, which portrays an interferogram intensity before adding noise. The strength of the noise which would be added to the interferogram in Figure 86 is depicted in Figure 87. The read noise variance is varied from 100 to 150,000 photons. At read noise levels below 100, the system is shot noise dominated while at read noise levels above 1000, the system is read noise dominated. Middle read noise levels between 100 and 1000 are a mixture, with high intensity subapertures being shot noise dominated while low intensity subapertures are read noise dominated.

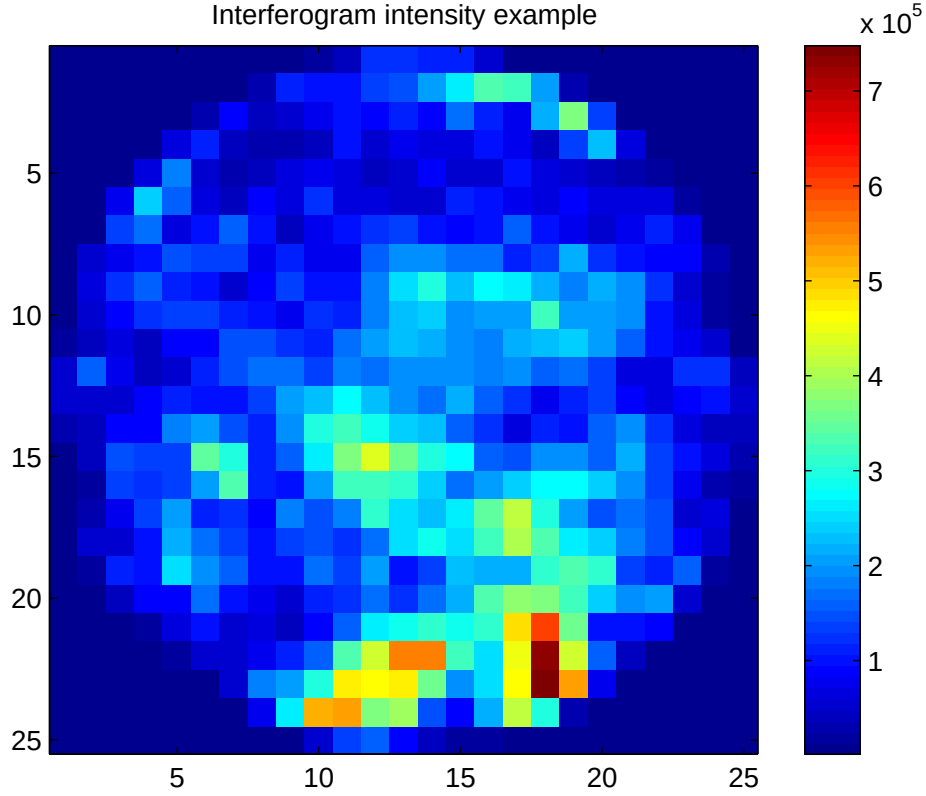


Figure 86: Interferogram intensity with colorbar scale

The performance of the system when varying read noise is portrayed in Figure 88. Qualitatively, the SRI is very robust to noise and the performance only degraded at the highest noise levels. The ‘knee’ in the performance curve depicted in Figure 88 is approximately 40,000, where the average SNR_M defined in Chapter II is less than 2.5.

A significant impact of higher noise levels on the system was in the difficulty in locking on the signal after closing the loop. This is indicated in Figure 89 which portrays the performance of a single realization for 100 frames at a read noise of 200,000. The system does not achieve steady-state until approximately 50 frames have passed, which is much higher than at lower noise level simulations. However, once the signal is acquired the system works well even at these much more significant noise levels.

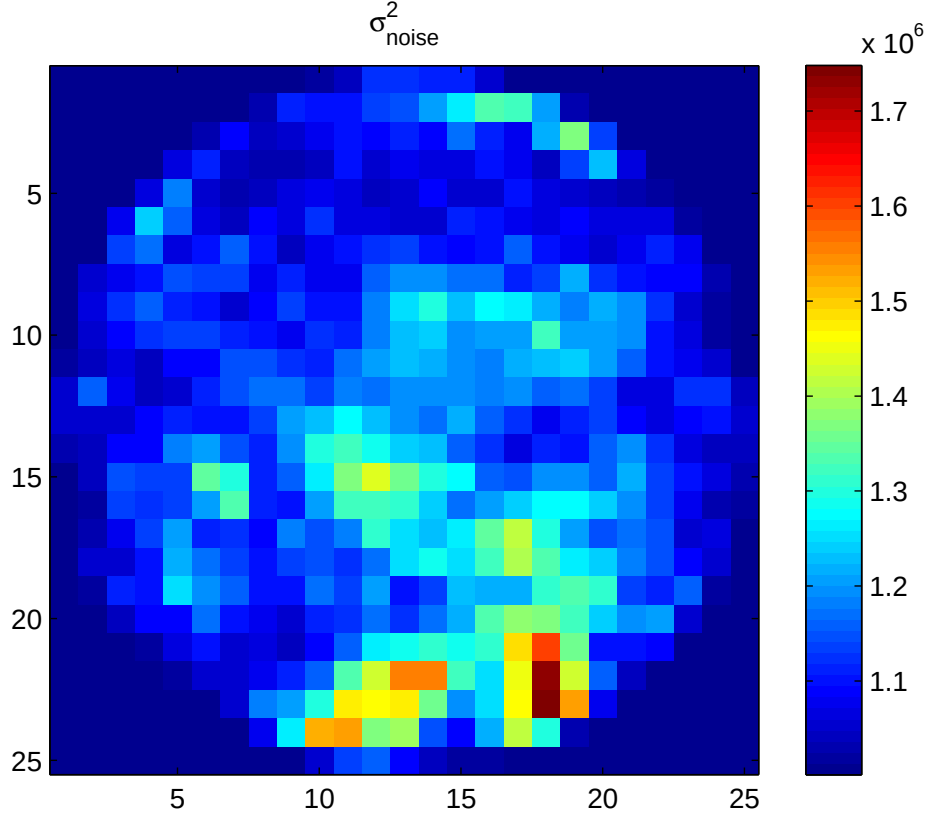


Figure 87: Interferogram noise variance with colorbar scale

6.6 Chapter conclusions

The purpose of this research is to design an AO system capable of effective operation in strong turbulence conditions. The system has been shown to operate well under these strong turbulence conditions for a reasonable range of atmospheric parameters other than simply having strong turbulence.

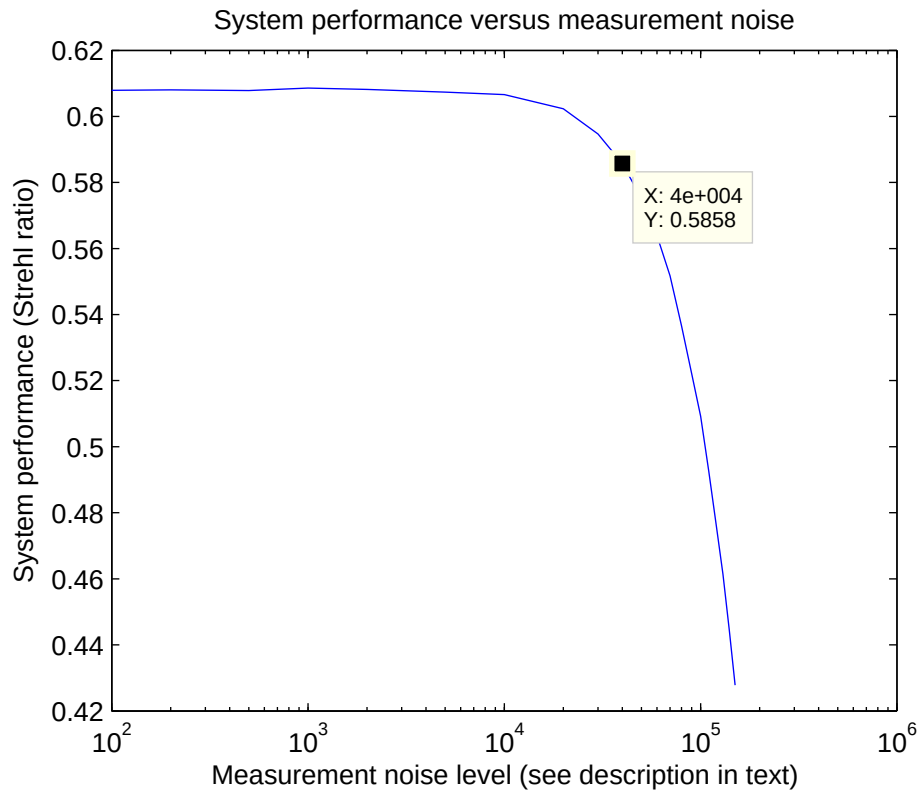


Figure 88: System performance versus measurement noise

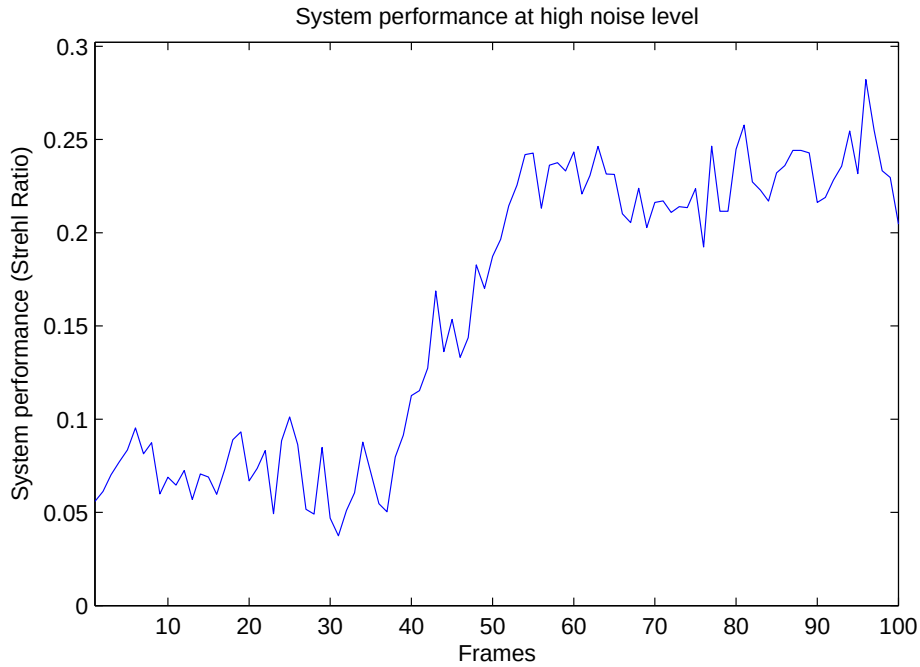


Figure 89: System performance at high noise levels depicting slow lock onto signal

VII. Conclusions

The development of an AO system effective under strong atmospheric turbulence has been a difficult problem which some consider the last and greatest challenge of AO. This research provides insight into the problem of AO control in strong turbulence which provides the basis for future AO control systems.

7.1 Significant Contributions

The following is a summary of the contributions resulting from this research.

7.1.1 Kalman estimation of anisoplanatic Zernike tilt. Published in the *Journal of Directed Energy*, this article utilized a Kalman filter to estimate the anisoplanatic tilt of a laser communication between a ground station and low-earth orbiting satellite. [43]

7.1.2 An improved temporally phase-shifted design. Presented at the SPIE ‘Optics and Photonics’ conference, this research developed an improved temporally phase-shifted SRI design which utilized an EKF to estimate the optical wavefront from individual SRI interferograms instead of utilizing four temporally-disparate interferograms. [41]

7.1.3 Recognition of the AO controller as an estimator. The recognition of the fact that a DM combined with a PI controller shares many of the same attributes as a Linear Kalman Filter justifies designs currently being used and allows targeted improvements to the system. Moreover, it justifies the common AO design structure used under weak turbulence. Specifically, a DM followed by a WFS feeding a PI controller provides adequate performance under both strong and weak atmospheric turbulence conditions. Recognizing this basic structure as capable, the development emphasis can be placed on dealing with WFSs, which are more effective under high scintillation, and the problems of unwrapping the phase of the optical field in the presence of branch points.

7.1.4 Unwrapping last. The difficulty in unwrapping the optical field is the crux of the problem in transitioning from systems designed to operate under weak turbulence to system designed to be able to handle strong turbulence. In fact, when considering that the commonly used AO system operating under weak turbulence utilizes a S-H WFS and

LS reconstructor which creates an unwrapped phase field, systems operating under weak turbulence conditions have not required unwrapping.

Under strong turbulence, unwrapping is much more important. When utilizing a non-gradient WFS such as the SRI, the output of the WFS is wrapped. Unwrapping the SRI output is insufficient, however, because the resultant phase field after the PI controller may be wrapped even if the error signal being integrated by the PI controller is unwrapped. Thus this work concludes that the optical field should be unwrapped *after* applying the PI control law instead of unwrapping the WFS output *before* the PI control law. Moreover, this research documents the effects of unwrapping at the wrong point in the AO design and identifies the cause of the performance degradations associated with simply unwrapping the SRI output.

7.1.5 Improved unwrapping. Unwrapping a wrapped field under weak turbulence results in a smooth phase field having adjacent phases within π of each other. With the exception of a constant offset (piston), this result is unique. In strong turbulence, an unwrapped field will still have discontinuities because of the branch cuts associated with the branch points of strong turbulence. The unwrapped field is then no longer unique. Even with a discrete field, the variations in placement of the branch cuts is significant. The problem becomes one of finding an effective placement of branch cuts which minimizes the impact of the branch cuts on system performance. This research developed a non-optimal but effective unwrapping procedure capable of unwrapping an optical field in the presence of branch points at a reasonable computational burden and is published in *Optics Express*. [42]

7.2 A single graph

The results of this research in developing a closed-loop AO system utilizing an SRI WFS can be portrayed in a single graph. The graph in Figure 90 depicts the performance of an AO system at the baseline parameters for a system using the approach advocated in this research of ‘control then unwrap’ (with ‘best-of-four’ unwrapping) against a system using ‘control then unwrap’ without ‘best-of-four’ unwrapping and finally a system utilizing ‘unwrap then control.’ The improvement is clearly seen by the superior Strehl ratio for the ‘Control then unwrap with improved unwrapper’ case.

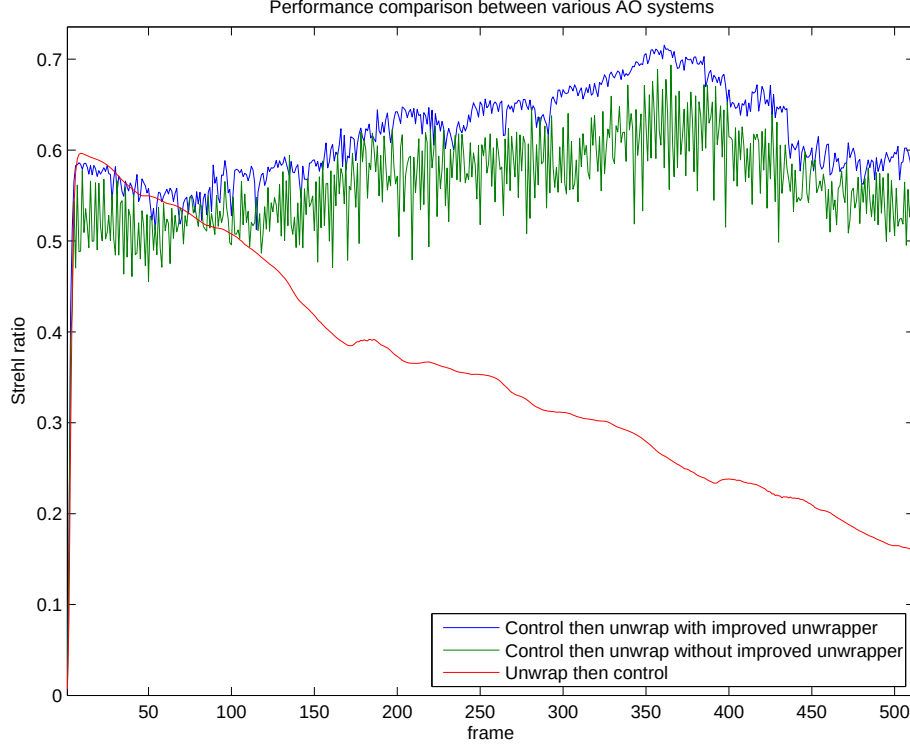


Figure 90: Comparison between AO systems

7.3 Future Work

Throughout this research, many areas worthy of further investigation were discovered. Maintaining a reasonable scope to this research, however, meant that they would have to be left for future endeavors. These areas are:

1. Expansion of the concepts in this research from simulation to the laboratory. The ASALT laboratory at SOR is an excellent platform to accomplish this task.
2. Laboratory testing of the improved unwrapper developed during this research.
3. Expansion of the system state-space to include $d\phi/dt$. This would potentially improve systems operating at lower sample rates.
4. Design of a system which has different Kalman gains for different areas of the aperture. An approximate SNR for the system could be determined based on signal strength. Areas with lower noise would have a higher Kalman gain to take advantage of the improved signal in those areas.

5. Incorporation of the spatial correlation between adjacent subapertures. This research concentrated solely on the temporal correlation of phase because that was deemed to be significantly stronger than spatial correlations under high sample rate conditions. Spatial correlation is also present and taking both temporal and spatial correlations into account has potential, particularly at lower sampling rates when temporal correlations weaken.
6. Investigation into designs which does not unwrap every frame. More specifically, designs which unwrap every other frame or every third frame in order to ease the computational burden associated with unwrapping. This is potentially useful if attempting to design systems which have a great many WFS subapertures and DM actuators but still need to operate very fast.
7. Investigation into systems which have WFSs which oversample the wavefront (have more subapertures than DM actuators). Unwrapping the DM last would still be the appropriate place to unwrap, but systems could be designed which could handle turbulence where the phase varied more than π between DM actuators.
8. Design the unwrapper to track the optimum phase range of the non-LS portion of the phase as a continuous variable instead of a discrete variable. Then, limit the amount of change that can take place in the range of the non-LS portion of the phase. Basically instead of choosing between four ranges at $\pi/2$ spacing, the optimal phase range would be tracked and the applied range would slowly vary instead of making $\pi/2$ jumps. This could potentially smooth out the noise of the demonstrated Strehl ratio performance. [27]

Appendix A. Appendix

This is where relevant code is. Unless otherwise stated, simulated data was generated using MATLAB.

A.1 Real and Imaginary contour generator code

```
%%%%%%%%%%%%%%
% This program takes a subset of a field and draws the contour slices where
% the real and imaginary parts of the field are zero.
% field2d is expected to already be loaded as a variable. 'start' and N
% (size of array segment) have to be chosen so that the segment of field2d
% looked at by the program does not exceed field2d
%%%%%%%%%%%%%%
clc;
% clear;
start_row=200; start_col=200; N=50;

% load 'I:\Classes\Wave Optics II\Project\Final1.mat' field1d
% [FieldPixSqrdd frames] = size(field1d);
% FieldPix = sqrt(FieldPixSqrdd);
% field2d=reshape(field1d,[FieldPix,FieldPix,frames]);
field=field2d(start_row:start_row+N-1,start_col:start_col+N-1,1);

% fid = fopen('L:\eng students\Cain\mantravadi\test fields\scon4r20_r');
% [field]=getufield(1,513,fid);
% fclose(fid);
% field=field(start_row:start_row+N-1,start_col:start_col+N-1,1);

% Isolate the section of field2d that will be looked at
R=sign(real(field)); I=sign(imag(field));

% Create 'real' horizontal and vertical lines. Envision horizontal and
```

```

% vertical lines between the NxN pixels. Thus there will be N-1 rows of
% N horizontal lines. Similarly there will be N rows of N-1 vertical lines.
% Create arrays of the real horizontal and real vertical lines - RH and RV
% respectively. Each point in RH and RV will be either a one or negative
% one. A one in RH indicates that the pixel above and below have the same
% polarity. Negative one indicates that the pixel above and below have
% opposite polarities. The RV matrix is similar except that it deals with
% pixels to the left and right instead of above and below.
RH=R(1:N-1,:).*R(2:N,:); % RH will be N-1 x N
RV=R(:,1:N-1).*R(:,2:N); % RV will be N x N-1

% Do the same thing for the imaginary portion of the field.
IH=I(1:N-1,:).*I(2:N,:); % IH will be N-1 x N
IV=I(:,1:N-1).*I(:,2:N); % IV will be N x N-1

% Draw out the figures. I couldn't figure out a way to do it except to
% isolate the negative ones in the RH, RV, IH and IV matrices, then
% determine their endpoints and plot them individually. Brute force, but
% adequate.
% figure(2)
% hold on
% RVZ=find(RV==-1);
% for ii=1:length(RVZ)
%     x = int16((RVZ(ii)-(N+1)/2)/N);
%     y = RVZ(ii)-x*N;
%     plot([x+1 x+1],[N+1-y N-y])
% end
% RHZ=find(RH==-1);
% for ii=1:length(RHZ)
%     x = int16((RHZ(ii)-N/2)/(N-1));
%     y = RHZ(ii)-x*(N-1);

```

```

%      plot([x x+1],[N-y N-y])
% end
% title('Real portion of field contour = 0 slice');
% legend('Real portion of field = 0')
%
% figure(3)
% hold on
% IVZ=find(IV== -1);
% for ii=1:length(IVZ)
%     x = int16((IVZ(ii)-(N+1)/2)/N);
%     y = IVZ(ii)-x*N;
%     plot([x+1 x+1],[N+1-y N-y],'r:')
% end
% IHZ=find(IH== -1);
% for ii=1:length(IHZ)
%     x = int16((IHZ(ii)-N/2)/(N-1));
%     y = IHZ(ii)-x*(N-1);
%     plot([x x+1],[N-y N-y],'r:')
% end
% title('Imaginary portion of field contour = 0 slice');
% legend('Imaginary portion of field = 0')

figure(4) hold on RVZ=find(RV== -1); RHZ=find(RH== -1);
IVZ=find(IV== -1); IHZ=find(IH== -1);

bp=bpfinder(angle(field)); pos_bps=find(bp==1);
neg_bps=find(bp== -1);

% Plot the first real line, imaginary line, positive and negative BPs. This properly se
% the legend
if length(RVZ)>0

```

```

        x = int32((RVZ(1)-(N+1)/2)/N);
        y = RVZ(1)-x*N;
        plot([x+1 x+1],[N+1-y N-y])
    end if length(IVZ)>0
        x = int32((IVZ(1)-(N+1)/2)/N);
        y = IVZ(1)-x*N;
        plot([x+1 x+1],[N+1-y N-y],'r:')
    end

    if length(pos_bps)>0
        x = int32((pos_bps(1)-N/2)/(N-1));
        y = pos_bps(1)-x*(N-1);
        plot(x+1,N-y,'o');
    end

    if length(neg_bps)>0
        x = int32((neg_bps(1)-N/2)/(N-1));
        y = neg_bps(1)-x*(N-1);
        plot(x+1,N-y,'x');
    end

    for ii=2:length(RVZ)
        x = int32((RVZ(ii)-(N+1)/2)/N);
        y = RVZ(ii)-x*N;
        plot([x+1 x+1],[N+1-y N-y])
    end

    for ii=1:length(RHZ)
        x = int32((RHZ(ii)-N/2)/(N-1));
        y = RHZ(ii)-x*(N-1);
        plot([x x+1],[N-y N-y])
    end

```

```

end

for ii=2:length(IVZ)
    x = int32((IVZ(ii)-(N+1)/2)/N);
    y = IVZ(ii)-x*N;
    plot([x+1 x+1],[N+1-y N-y],'r:')
end

for ii=1:length(IHZ)
    x = int32((IHZ(ii)-N/2)/(N-1));
    y = IHZ(ii)-x*(N-1);
    plot([x x+1],[N-y N-y],'r:')
end

title('Real and Imaginary portions of field contours = 0 slices
(overlay)'); if length(pos_bps)>0 & length(neg_bps)>0
    legend('Real portion of field = 0','Imaginary portion of field = 0', 'Positive Branch')
else
    if length(pos_bps)>0
        legend('Real portion of field = 0','Imaginary portion of field = 0', 'Positive Branch')
    else
        if length(neg_bps)>0
            legend('Real portion of field = 0','Imaginary portion of field = 0', 'Negative Branch')
        else
            legend('Real portion of field = 0','Imaginary portion of field = 0')
        end
    end
end

end

bp=bpfinder(angle(field)); pos_bps=find(bp==1);
neg_bps=find(bp==-1);

```

```

for ii=2:length(pos_bps)
    x = int32((pos_bps(ii)-N/2)/(N-1));
    y = pos_bps(ii)-x*(N-1);
    plot(x+1,N-y,'o');
end

```

```

for ii=2:length(neg_bps)
    x = int32((neg_bps(ii)-N/2)/(N-1));
    y = neg_bps(ii)-x*(N-1);
    plot(x+1,N-y,'x');
end

```

```

axis([0 N 0 N]);

```

A.2 Branch point finder

```

% bpfinder find branch points in a phase field
% [bp]=bpfinder(pfield)
% takes a 2 dimensional phase front and returns a 2-dimensional image
% with 1's and -1's corresponding to branch points in that phase front
% Originally made by Jai Montravadi, modified by Todd Venema

```

```

function [bp]=bpfinder(pfield)

d1=diff(pfield,1,1); d2=diff(pfield,1,2);
d1=(d1<-pi).*2.*pi-(d1>pi).*2.*pi+d1;
d2=(d2<-pi).*2.*pi-(d2>pi).*2.*pi+d2;
bp=[d1(:,2:end)-d2(2:end,:)-d1(:,1:end-1)+d2(1:end-1,:)]; bp =
(bp>0.1)-(bp<-0.1);

return

```

A.3 pdf maker

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
% This code generates a pdf of the error from an SRI WFS. It does so by  
% generating 1000^2 samples. Each sample has four interferograms and then  
% noise is added to each interferogram. The interferograms are then used  
% to generate phases via the arctan function. Errors are determined as  
% the difference between the estimate and truth. The errors are then put  
% into a histogram and the histogram is plotted. This histogram is  
% effectively an empirical pdf of the error signal.  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
clear;
```

```
clc;
```

```
noise_var=0.3;
```

```
samples=1000;
```

```
real_phi=pi/4;
```

```
photons_per_unit=60;
```

```
A=1;
```

```
Ar=1;
```

```
x=A*cos(real_phi);
```

```
y=A*sin(real_phi);
```

```
M=(Ar+x).^2+y.^2;
```

```
M1=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
```

```
SNR1=M/(noise_var+M/photons_per_unit)
```

```
M=x.^2+(Ar+y).^2;
```

```
M2=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
```

```
SNR2=M/(noise_var+M/photons_per_unit)
```

```
M=(Ar-x).^2+y.^2;
```

```
M3=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
```

```

SNR3=M/(noise_var+M/photons_per_unit)
M=x.^2+(Ar-y).^2;
M4=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
SNR4=M/(noise_var+M/photons_per_unit)

phi=mod(atan2(M2-M4,M1-M3)-real_phi+pi,2*pi)-pi;
[y x]=hist(phi(:),100);
y=y./sum(y)*99/(x(100)-x(1));
y2=exp(-x.^2./2./std(phi(:)).^2);
y2=y2./sum(y2).*99./(x(100)-x(1));
plot(x,y,x,y2)
axis([-1 1 0 2.5])
legend('pdf of phi_{SRI}','Gaussian pdf w/ same variance')
xlabel('\phi_{error}')
ylabel('pdf')
set(gcf,'Position',[50 400 800 400])
set(gcf,'PaperPositionMode','auto')
% print -depsc I:\Dissertation\CH3_Methodology\figures\pdf

```

A.4 Error variance for different phases

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This code generates a plot of error variance as phase changes for an
% SRI WFS. It works by generating 200^2 samples of interferograms for
% 360 different angles for a given signal amplitude. The variance at
% each angle is determined and the different variances are plotted
% against their phases.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear;
clc;

```

```

noise_var=.3;
samples=300;
photons_per_unit=60;
A=1;
Ar=1;
real_phi=zeros(360,1);
SD=zeros(360,1);

for ii=1:360
    ii
    real_phi(ii)=(ii-1)*2*pi/360;
    x=A*cos(real_phi(ii));
    y=A*sin(real_phi(ii));
    M=(Ar+x).^2+y.^2;
    M1=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
    M=x.^2+(Ar+y).^2;
    M2=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
    M=(Ar-x).^2+y.^2;
    M3=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);
    M=x.^2+(Ar-y).^2;
    M4=M+(noise_var+M/photons_per_unit).^0.5.*randn(samples);

    est_phi_error=mod(atan2(M2-M4,M1-M3)-real_phi(ii)+pi,2*pi)-pi;
    SD(ii)=std(est_phi_error(:));
end

plot(real_phi,SD)
axis([0 2*pi 0.1 1])
xlabel('\phi')
ylabel('\sigma')
set(gcf,'Position',[50 400 800 400])

```

```
set(gcf,'PaperPositionMode','auto')
% print -depsc I:\Dissertation\CH3_Methodology\figures\st_dev_vs_phi
```

A.5 Variance generator

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This code generates a surface of noise variance for an SRI WFS. It
% works by generating 200^2 samples of interferograms for a given
% refererence and signal amplitude (Ar and A). The interferograms have
% noise of a given strength added to them and the the estimated phase is
% computed from the standard arctan function. The errors are determined
% difference between the estimate and truth. The errors are put into a
% histogram which is effectively an empirically determined pdf of the
% error. The standard deviation of the pdf then plotted against Ar and A
% is stored to show how they affect the error variance.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear;
clc;

read_noise=50000;
samples=200;
real_phi=pi/4;
span=40;
actual_SD=zeros(span);
est_SD=zeros(span);
Ar=ones(span,1);
As=ones(span,1);
mx_Ar=1300;
mx_As=260;

for jj=1:span
```

```

    As(jj)=jj*mx_As/span;
%    jj
    for kk=1:span
        Ar(kk)=kk*mx_Ar/span;
        x=As(jj)*cos(real_phi);
        y=As(jj)*sin(real_phi);

        M=(Ar(kk)+x).^2+y.^2;
        M1=M+sqrt(read_noise.^2+M).*randn(samples);
        M=x.^2+(Ar(kk)+y).^2;
        M2=M+sqrt(read_noise.^2+M).*randn(samples);
        M=(Ar(kk)-x).^2+y.^2;
        M3=M+sqrt(read_noise.^2+M).*randn(samples);
        M=x.^2+(Ar(kk)-y).^2;
        M4=M+sqrt(read_noise.^2+M).*randn(samples);

        phi=mod(atan2(M2-M4,M1-M3)-real_phi+pi,2*pi)-pi;

        actual_var(jj,kk)=var(phi(:));
        est_var(jj,kk)=3.29.*(exp(-As(jj).*Ar(kk)./35000)).^2;
    end
end

figure(1)
surf(Ar,As,actual_var)
ylabel('A_R')
xlabel('A_S')
zlabel('\sigma^2_\phi')
title('\sigma^2_\phi from Monte Carlo analysis')
camorbit(75,-24)
% set(gcf,'Position',[50 400 800 400])

```

```

% set(gcf,'PaperPositionMode','auto')
% print -depsc I:\Dissertation\CH_A0_as_estimator\figures\Monte_Carlo_variance

figure(2);
surf(Ar,As,est_var);
ylabel('A_R')
xlabel('A_S')
zlabel('\sigma^2_\phi')
title('estimated \sigma^2_\phi')
camorbit(75,-24)
% set(gcf,'Position',[50 400 800 400])
% set(gcf,'PaperPositionMode','auto')
% print -depsc I:\Dissertation\CH_A0_as_estimator\figures\estimated_variance

adj=mean(mean(est_var-actual_var))

figure(3)
surf(Ar,As,est_var-actual_var)

sum(sum(abs(actual_var-est_var)))

Qd=0.0583;
As2d= repmat(As,1,span);
Ar2d= repmat(Ar',span,1);
RoverQd=3.29.*(exp(-As2d.*Ar2d./17500)).^1./Qd;

K=1./(2.*RoverQd).*((1+4.*RoverQd).^0.5-1);
figure(4)
surf(Ar,As,K)
title('K vs. A_R and A_S')
ylabel('A_R')

```

```
xlabel('A_S')
xlabel('K')
% set(gcf,'Position',[50 400 600 400])
% set(gcf,'PaperPositionMode','auto')
% print -depsc I:\Dissertation\CH_AO_as_estimator\figures\K_vs_AR_AS
```

Bibliography

1. Andrews, Larry C. and Ronald C. Philips. *Laser Beam Propagation through Random Media*. SPIE Optical Engineering Press, 1998.
2. Apostol, Tom M. *Mathematical Analysis*. Addison Wesley Longman, second edition, 1974.
3. Barchers, Jeffrey. "The performance of wavefront sensors in strong turbulence". *Proc. SPIE*, 4839:217–227, Apr 2003.
4. Barchers, Jeffrey. "Personal Communication", Dec 2007.
5. Barchers, Jeffrey. "Personal Communication", Jun 2008.
6. Brennan, Terry. "Personal Communication", Oct 2007.
7. Brennan, Terry J. and Phillip H. Roberts. "AOTools The Adaptive Optics Toolbox-Users Guide version 1.2".
8. Byron M. Welsh, Michael C. Roggemann, Brent L. Ellerbroek and Timothy L. Pennington. "Fundamental performance comparison of a Hartmann and a shearing wave-front sensor". *Appl. Opt.*, 34(21):4186–4195, July 1995.
9. Cheng, David K. *Fundamentals of Engineering Electromagnetics*. Prentice Hall, first edition, 1992.
10. Corley, Melissa S. and Troy Rhoadarmer. "Evaluation of phase-shifting techniques for a self-referencing interferometer wavefront sensor". *Proc. SPIE*, 5894(0R), 2005.
11. David Dayton, Bob Pierson and Brian Spielbusch. "Atmospheric structure gunction measurements with a Shack-Hartmann wave-front sensor". *Optics Letters*, 17(24):1737–1739, December 1992.
12. David Gerwe, James Stone and Harold Schall. "Closed-Loop Control for Adaptive Optics Wavefront Compensation in Highly Scintillated Condtions". *Unknown*.
13. Dereniak, E.L. and G.D. Boreman. *Infrared Detectors and Systems*. John Wiley & Sons, Inc., 1996.
14. Fried, David L. "Branch point problem in adaptive optics". *J. Opt. Soc. Am. A*, 15(10):2759–2768, October 1998.
15. Fried, David L. "The Nature of the Branch Point Problem in Adaptive Optics". *Proc. SPIE*, 3381, Apr 1998.
16. Fried, David L. "Adaptive optics wave function reconstruction and phase unwrapping when branch points are present". *Optics Communications*, 200(1-6):43–72, December 2001.
17. Fried, David L. and Jeffrey L. Vaughn. "Branch cuts in the phase function". *Appl. Opt.*, 31(15):2865–2882, May 1992.
18. Ghiglia, Dennis C. and Mark D. Pritt. *Two-Dimensional Phase Unwrapping Theory, Algorithms, and Software*. John Wiley & Sons, Inc., 1998.

19. Goodman, Joseph W. *Statistical Optics*. Wiley Interscience, 1985.
20. Goodman, Joseph W. *Introduction to Fourier Optics*. McGraw-Hill, second edition, 1996.
21. Hardy, John W. *Adaptive Optics for Astronomical Telescopes*. Oxford University Press, 1998.
22. Hecht, Eugene. *Optics*. Addison Wesley, fourth edition, 2002.
23. Horowitz, Paul and Winfield Hill. *The Art of Electronics*. Cambridge University Press, second edition, 2002.
24. Klein, Laura and Troy Rhoadarmer. "Performance measurements of a self-referencing interferometer wavefront sensor with optical amplification". *Proc. SPIE*, 5894(0Q), 2005.
25. Klein, Laura and Troy Rhoadarmer. "Closed-loop control techniques for an adaptive-optical system with an interferometric wavefront sensor". *Proc. SPIE*, 6306, 2006.
26. Krane, Kenneth. *Modern Physics*. John Wiley & Sons, Inc., 1983.
27. Mantravadi, Jai. "Personal Communication", Jun 2008.
28. Maybeck, Peter S. *Stochastic Models, Estimation, and Control Volume 1*. Navtech Book & Software Store, 1994.
29. Maybeck, Peter S. *Stochastic Models, Estimation, and Control Volume 2*. Navtech Book & Software Store, 1994.
30. Maybeck, Peter S. *Stochastic Models, Estimation, and Control Volume 3*. Navtech Book & Software Store, 1994.
31. Ogata, Katsuhiko. *Modern Control Engineering*. Prentice-Hall, third edition, 1997.
32. Primmerman, Charles A. "Atmospheric-compensation experiments in strong-scintillation conditions". *Appl. Opt.*, 34(12):2081–2088, Apr 1995.
33. Reagan, Ronald. "'Star Wars' speech, 23 March 1983".
34. Rhoadarmer, Troy. "Slide from presentation presented by Dr Rhoadarmer to AFIT in Summer 2006. Titles updated for clarity."
35. Rhoadarmer, Troy. "Development of a self-referencing interferometer wavefront sensor". *Proc. SPIE*, 5553:112–126, October 2004.
36. Rhoadarmer, Troy. "Personal Communication", March 2007.
37. Rhoadarmer, Troy and Jeffrey Barchers. "Noise analysis for complex field estimation using a self-referencing interferometer wave front sensor". *Proc. SPIE*, 4825:215–227, 2002.
38. Rhoadarmer, Troy and Laura Klein. "Design of a spatially phase shifted self-referencing interferometer wave front sensor". *Proc. SPIE*, 6306, 2006.
39. Roggemann, Michael C. "Branch-point reconstruction in laser beam projection through turbulence with finite-degree-of-freedom phase-only wavefront correction". *J. Opt. Soc. Am. A*, 17(1):53–62, Jan 2000.

- 40. Schmidt, Jason. “Figure provided by Dr. Jason Schmidt”.
- 41. Venema, Todd M. “An improved temporally phase-shifted design”. *SPIE (conference proceedings paper)*, 6711, Sep 2007.
- 42. Venema, Todd M. and Jason Schmidt. “Optical phase unwrapping in the presence of branch points”. *Optics Express*, 16(10):6985–6998, May 2008.
- 43. Venema, Todd M. and Juan R. Vasquez. “Kalman estimation of anisoplanatic Zernike tilt”. *Journal of Directed Energy*, 2(4), Fall 2007.
- 44. Verdeyen, Joseph T. *Laser Electronics*. Prentice-Hall, third edition, 1995.
- 45. Website, SOR. URL <http://www.de.afrl.af.mil/SOR/index.htm>. Figure and text obtained from SOR Website by Dr Jason Schmidt, date unknown.
- 46. Wikipedia. March 2007. URL <http://en.wikipedia.org/wiki/Vector-potential>. Website.

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.						
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE			3. DATES COVERED (From — To)	
11-09-2008		Doctoral Dissertation			May 2007 — Sep 2008	
4. TITLE AND SUBTITLE Closed-Loop Adaptive Optics Control in Strong Atmospheric Turbulence				5a. CONTRACT NUMBER		
				N/A		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Todd M. Venema, LtCol, USAF				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765					8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/DEE/ENG/08-21	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) None					10. SPONSOR/MONITOR'S ACRONYM(S)	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approval for public release; distribution is unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT A self-referencing interferometer based closed-loop adaptive optics controller is developed which is designed to operate effectively under strong turbulence conditions. The aberrated optical field is modeled stochastically and then estimates of the state of the system are developed using a steady-state, fixed-gain Kalman filter. The phase of the optical field is considered the state of the system which is wrapped in a limited range of $(-\pi, \pi]$. This phase is unwrapped through the use of a least-squares reconstructor which has been modified to work effectively in the presence of branch points associated with strong turbulence. The conjugate of the optical phase is then applied to the system's deformable mirror in order to correct for the effects of atmospheric turbulence on the optical field. The advances developed in this research are in the application of a steady-state, fixed-gain Kalman filter to the input of an adaptive optic system, unwrapping the optical phases after the field estimation, and improving the phase unwrapping by varying the domain of the rotational phase component present in strong turbulence. The system developed in this research is shown in computer simulation to be improved over current designs by comparing performance plots of system Strehl ratios for systems utilizing the different designs.						
15. SUBJECT TERMS Adaptive Optics, Strong Turbulence, Self-Referencing Interferometers, Kalman Filters						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Juan Vasquez	
U	U	U	UU	150	19b. TELEPHONE NUMBER (include area code) (937) 427-9725 / juan.vasquez@numerica.us	