Comparison of the Performance of Modal Control Schemes for an Adaptive Optics System and Analysis of the Effect of Actuator Limitations

Masaki Nagashima, Jae Jun Kim, Brij Agrawal

Abstract

In this study, modal control with various bases has been experimentally evaluated on an experimental Adaptive Optics (AO) testbed. In most custom-built adaptive optics control problems, spatial resolution and available stroke of the deformable mirror (DM) are assumed to be sufficient for the intended performance. In practice, however, there are situations where DMs of limited capability are used and the control system is desired to make most out of the hardware even in such situations. In this paper, the effect of the hardware limitations, such as the spatial resolution of the wavefront sensor and the deformable mirror (DM) on the performance of the AO system, is investigated and several control techniques are proposed to mitigate the problem. The derivatives of Zernike polynomials do not have the orthogonal property.A simple approximate orthogonalization technique is proposed along with the method to avoid the discrepancy of minimum error in the sensor output vector space and in the vector space where the control law is applied. An anti-windup technique is also proposed to reduce the adverse effect by an integral controller when saturation of the DM occurs. Effectiveness of the proposed techniques is demonstrated by experiments on the pseudo real-time AO testbed.

Keywords: adaptive optics, adaptive filters, modal reduction, actuator saturation.

I. INTRODUCTION

Adaptive Optics (AO) refers to an optical control system used in telescope or laser propagation systems where the phase aberration of the incoming beam is measured by a wavefront sensor (WFS) and the incoming beam, or the outgoing beam sharing the same optical path, is compensated by a Deformable Mirror (DM). Feedback control has been often used to address dynamic aberration such as atmospheric disturbance, but in recent years, more advanced control schemes such as optimal control and adaptive control have been proposed [1–4].

Adaptive optics is often modelled as a statically coupled Multi-Input Multi-Output (MIMO) problem in which the dynamic coupling is ignored based on the assumption that the response of the deformable mirror is fast compared with the sample rate of the system. The plant is then represented by a matrix and pure step delays, and the control problem can be addressed by so-called modal control where the error is projected onto a basis to apply a control law in a vector space that is different from that of the sensor output. Any basis can be used, but they have different effects on the performance.

In this paper, performances of modal control with various bases have been evaluated experimentally on an AO testbed, with a focus on the effect of hardware limitations that stem from the spatial resolution and the finite stroke of the DM. Two control techniques, namely, modified projection method and anti-windup control are

proposed to mitigate the effect of the DM limitations.A simple method to orthogonalize Zernike derivative basis is proposed.

For an AO system where the number of DM actuator channels is less than the number of WFS measurement points, the control system is over-determined and one can only obtain a solution that minimizes the error according to a certain definition such as the lest square error solution. The error by such a solution is the physical lower bound of the error for the given system, and the objective of the control system is to achieve this minimum error. However, an arbitrary basis does not necessarily produce this minimum error with modal control even when the basis vectors are orthogonal. The proposed modified projection method remedies this problem by removing the uncontrollable subspace components from the measured error.

When saturation of the DM occurs due to a temporarily surge of the error, the state of the integral controller tends to "wind-up" and it takes a long time for the state to go back to the level before saturation. Even after the error level returns to the level that the controller does not need to cause saturation to compensate it. This problem is inherent in this type of controller and an application of a so-called anti-windup technique is proposed which can mitigate the symptom. The anti-windup controller computes the difference between the actual plant response and the computed plant response without saturation, which is the effect of the controller eliminated due to the saturation. By adding this estimated controller effect to the error fed to the controller, the anti-windup controller can avoid the overcorrection by the controller to prevent windup. The anti-windup control cannot prevent actuator saturation, but it can reduce the adverse effect of the controller when saturation occurs.

Derivatives of Zernike polynomials evaluated at discrete points used as the basis for Zernike decomposition are not orthogonal. A simple method based on Singular Value Decomposition (SVD) is proposed which orthogonalizes the Zernike derivative basis while preserving the spatial characteristics of the Zernike polynomials to some extent at lower orders.

The performances of the proposed methods are experimentally evaluated with a Proportional-Integral (PI) controller and a PI controller augmented using a single channel filtered-x (FX) Recursive Least Squares (RLS) adaptive filter control with three commonly used bases and the proposed basis, namely, the influence matrix (the pseudo inverse method), the Singular Value Decomposition (SVD) basis, the Zernike derivative basis, and the proposed orthogonal Zernike derivative basis.

This paper is organized as follows: In the next section, modal control and various modal bases considered in this paper, as well as the FXRLS adaptive filter control, are briefly reviewed. Section III presents the proposed control techniques, and Section IV describes the AO testbed used for experiments and presents the experimental results. The conclusion and discussions are given in Section V.

II. MODAL CONTROL OF ADAPTIVE OPTICS SYSTEM

Figure 1 shows a model of the AO control system considered in this paper. The controller input $\mathbf{e}(k)$ and output $\mathbf{u}(k)$, which are the wavefront error and the DM command, respectively, are expressed as vectors. The dynamics of the DM are ignored and the path from $\mathbf{u}(k)$ to $\mathbf{y}(k)$ is represented by a constant matrix $\mathbf{\Gamma}$ called the poke or influence matrix. Any delays in the system due to WFS and DM hardware operations are separated from the poke matrix and represented by a step delay z^{-q}. The disturbance represented by vector φ is the phase aberration measured by the WFS. For Shack-

Hartmann (SH) WFS, the sensor output is the local slope of the phase and the number of the output channel is twice the number of the measurement points. Each column vector of the poke matrix represents the steady state response of the wavefront sensor for a unit input to a single channel of the DM. Since the poke matrix is usually not a diagonal matrix, i.e., there are static couplings of the channels, the system is a MIMO system.

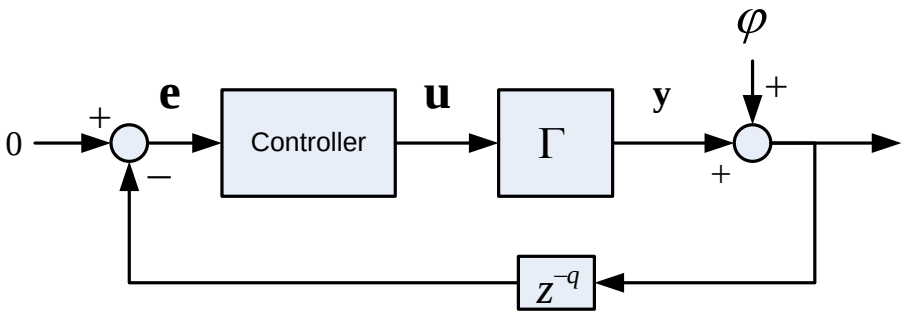


Figure 1 AO system model

One of the simplest approaches to design a controller for this MIMO system is decoupling the input and output channels, i.e., diagonalizing the poke matrix. One of the advantages of this approach is that a Single Input Single Output (SISO) controller can directly be applied whose computational cost is usually lower than that of a MIMO controller. One of the disadvantages is that it is not always possible to obtain a very accurate model of the plant and the performance and stability of the controller can suffer if the modelling error is large.

When $\mathbf{\Gamma}$ is an invertible square matrix, decoupling of the system can be done by multiplying the sensor output with the inverse of the poke matrix to make the matrix relating $\mathbf{u}(k)$ to $\mathbf{e}(k)$ an identity matrix. It is often the case, however, that $\mathbf{\Gamma}$ is not a square matrix or does not have the full rank. In such a case, decoupling of each control channel can still be achieved by projecting the sensor output to a different vector space referred to here as the control space. Figure 2 shows a diagram of this type of control method know as modal control. The matrix **F** transforms the sensor output to the control space and the matrix **G** transforms it to the actuator command.

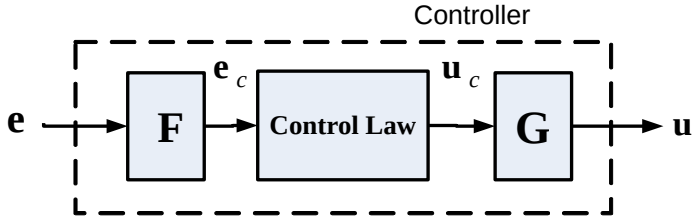


Figure 2 Controller structure for modal control

For $\mathbf{\Gamma} \in \Re^{m \times n}$ whose rank is r, there exists a pair of matrices $\mathbf{F} \in \Re^{r \times m}$ and $\mathbf{G} \in \Re^{n \times r}$ that can uncouple the path from $\mathbf{u}_c$ to $\mathbf{e}_c$. The matrix **F** can be constructed by choosing the row vectors from any set of linearly independent vectors in $\Re^m$ that spans the column vector space of $\mathbf{\Gamma}$. The matrix $\mathbf{F\Gamma}$ is then full rank and choosing $\mathbf{G} = (\mathbf{F\Gamma})^{\dagger}$

will diagonalize the open loop path $\mathbf{F\Gamma}$, where † denotes the pseudo inverse which is defined as follows for a full rank matrix $\mathbf{A} \in \Re^{m \times r}, r \le m$.

$$\mathbf{A}^{\dagger} = \left(\mathbf{A}\mathbf{A}^{T}\right)^{-1}\mathbf{A}^{T} \tag{1}$$

Note that the number of row vectors of **F** can be reduced arbitrarily without affecting the decoupling of the system. This leads to the idea of modal reduction where the controller space is restricted to a subspace of the poke matrix column vector space in order to reduce the computation and increase the robustness. Once decoupling of the open-loop path is established, the feedback system can be treated as a set of SISO feedback loops and a single SISO control law can be applied independently to each channel.

Modal Basis

In this paper, three common bases are investigated as the basis for modal control, namely column vectors of the poke matrix, the Singular Value Decomposition (SVD) basis, and the Zernike derivative basis. These bases are described in this section.

Poke Matrix

If $\mathbf{\Gamma} \in \Re^{m \times n}$ is full rank and $n \le m$, then $\mathbf{F} = \mathbf{\Gamma}^{\dagger}$ and $\mathbf{G} = \left(\mathbf{\Gamma}^{\dagger}\mathbf{\Gamma}\right)^{\dagger} = \mathbf{I}$ can uncouple the open loop path and it is equivalent to the conventional pseudo inverse poke matrix approach. Here, **I** is an identity matrix of the appropriate dimension. In this case, the SISO control law C(z) is applied to the error converted by $\mathbf{\Gamma}^{\dagger}$ as follows:

$$\mathbf{u}(z) \quad = C(\mathrm{z})\mathbf{\Gamma}^{\dagger}\mathbf{e}_s(z) = C(\mathrm{z})\left(\mathbf{\Gamma}^{T}\mathbf{\Gamma}\right)^{-1}\mathbf{\Gamma}^{T}\mathbf{e}_s(z) = \left(\mathbf{\Gamma}^{T}\mathbf{\Gamma}\right)^{-1}\left(C(\mathrm{z})\mathbf{I}\right)\mathbf{\Gamma}^{T}\mathbf{e}_s \tag{2}$$

It can be seen that it is equivalent to applying the control law to the sensor error projected onto the column vectors of the poke matrix and that $\mathbf{F} = \mathbf{\Gamma}^{T}$ and $\mathbf{G} = (\mathbf{\Gamma}^{T}\mathbf{\Gamma})^{-1}$.

The column vectors of the poke matrix, referred to as the poke matrix basis hereafter, represent the slope of the phase generated by the DM when unit input is applied to each channel separately. The poke matrix basis of the DM - WFS system used in the experiment is obtained from the measurement and the phase of the slope vectors (column vectors) is reconstructed by a zonal method proposed by Southwell [5] with a slight modification for the hexagonal geometry of the WFS. Figure 3 shows the first 36 basis vectors.

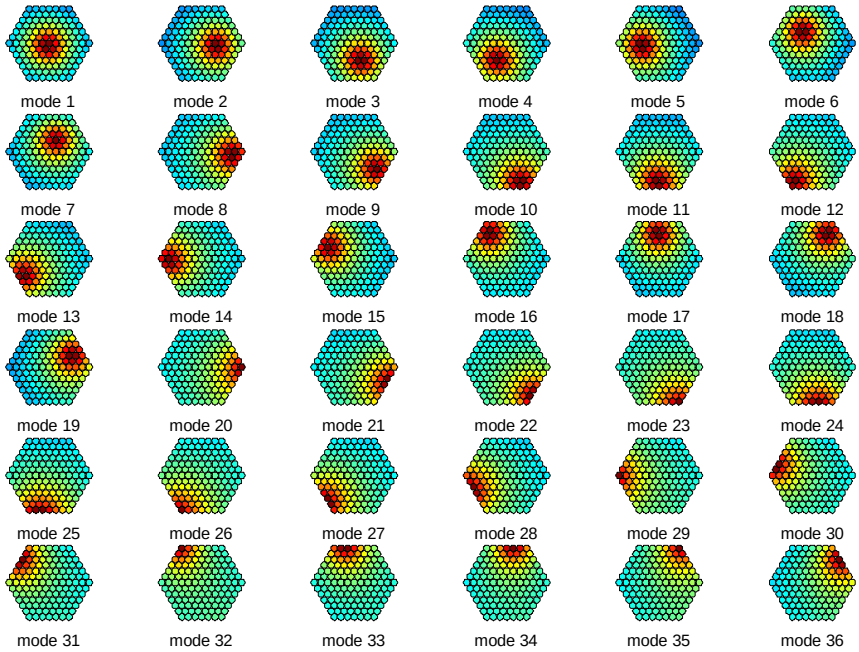


Figure 3 Visualization of Poke matrix basis

Singular Value Decomposition

The poke matrix can be decomposed into three orthogonal matrices by the singular value decomposition as follows:

$$\mathbf{\Gamma} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{3}$$

Here, $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices whose column vectors are orthogonal and normalized. For $\mathbf{\Gamma} \in \Re^{m\times n}$, the size of $\mathbf{U}$ and $\mathbf{V}$ are $m \times m$ and $n \times n$, respectively. The matrix $\mathbf{\Sigma}$ is a diagonal matrix whose diagonal components are the singular values of the matrix $\mathbf{\Gamma}$. For a poke matrix whose rank is $r < n < m$, only the first r diagonal components of $\mathbf{\Sigma}$ have non-zero values and the dimension of $\mathbf{\Sigma}$ can be reduced to $r \times r$ with $\mathbf{U}$ and $\mathbf{V}$ also reduced to have only first r columns, which are denoted as $\mathbf{\Sigma}' \in \Re^{r\times r}$, $\mathbf{U}' \in \Re^{m\times r}$, and $\mathbf{V}' \in \Re^{n\times r}$. Control law can be applied to the error projected onto the column vectors of $\mathbf{U}'$, and the conversion to the actuator command can be obtained by:

$$\mathbf{u} = \left(\mathbf{U}'^T\mathbf{\Gamma}\right)^{\dagger}\mathbf{u}_c = \mathbf{V}^T\left(\mathbf{\Sigma}'\right)^{-1}\mathbf{u}_c \tag{4}$$

Figure 4 is the visualization of the phase reconstructed from the SVD basis vectors in $\mathbf{U}'$ for the first 36 modes.

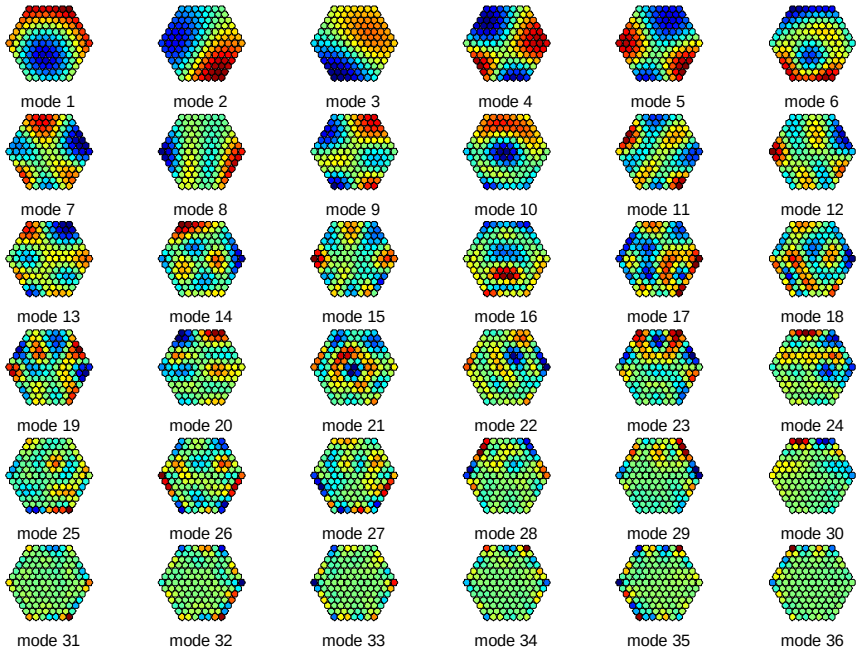


Figure 4 Visualization of SVD basis

Zernike Basis

Zernike basis vectors are constructed from the Zernike polynomial by evaluating the Zernike polynomial at the coordinates of the sensor measurement points normalized by the diameter length. Zernike polynomial is a sequence of polynomials that are orthogonal and normal over the interior of a unit circle and commonly used in adaptive optics and optics in general because of its convenience for circular aperture optical components.

The series of the fringe Zernike polynomials is given by

$$Z_n^m(\rho,\theta) = R_n^m(\rho)e^{im\theta} \tag{5}$$

$$R_n^m(\rho) = \sum_{s=0}^{\frac{n-m}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)} \rho^{n-2s} \tag{6}$$

where R_n^m represents the radial function of a normalized radius $\rho = r/R$ and R and r are the radius of the aperture and the radius of a point in the aperture, respectively. The indices n and m are both integers with $n \geq m$, $n \geq 0$. There are variations of the polynomial for different applications, and derivations of the standard form of Zernike polynomials can be found in [5], [6]. The aberrations described by the lower order Zernike polynomials are commonly referred to by the names, such as focus, coma, astigmatism, and etc. Zernike polynomials can be transformed into Cartesian coordinates by the relationship between the polar and Cartesian coordinates.

The coefficient of the Zernike polynomial can be obtained from the slope measurement of the SH WFS by the partial derivatives of Zernike polynomials as follows [7]:

$$\mathbf{c} = d\mathbf{Z}^{\dagger}\mathbf{e}_{\mathbf{s}} \tag{7}$$

where $\mathbf{c}$ is the Zernike coefficient, $\mathbf{e}_{\mathbf{s}}$ is the slope measurement, and the matrix $d\mathbf{Z}^{\dagger}$ is the pseudo inverse of $d\mathbf{Z}$ whose column vectors are the partial derivative polynomial evaluated at the coordinates points written as follows:

$$d\mathbf{Z} = \left[\frac{\partial \mathbf{Z}(x,y)}{\partial x} \quad \frac{\partial \mathbf{Z}(x,y)}{\partial y} \right]^{T} \tag{8}$$

The pseudo inverse $d\mathbf{Z}^{\dagger}$ is defined as:

$$d\mathbf{Z}^{\dagger} = \left(d\mathbf{Z}^{T} d\mathbf{Z} \right)^{-1} d\mathbf{Z}^{T} . \tag{9}$$

The size of $d\mathbf{Z}$ is $2m \times n$, where m is the number of sensor spots and n is the number of Zernike modes to be included. If $d\mathbf{Z}^{\dagger}\mathbf{\Gamma}$ is full rank, matrices $\mathbf{F} = d\mathbf{Z}^{\dagger}$ and $\mathbf{G} = \left(d\mathbf{Z}^{\dagger}\mathbf{\Gamma} \right)^{\dagger}$ decouple the open loop . Figure 5 is the visualization of the phase reconstructed from the partial derivative vectors in $d\mathbf{Z}$ for the first 36 modes obtained for the WFS geometry considered in this paper.

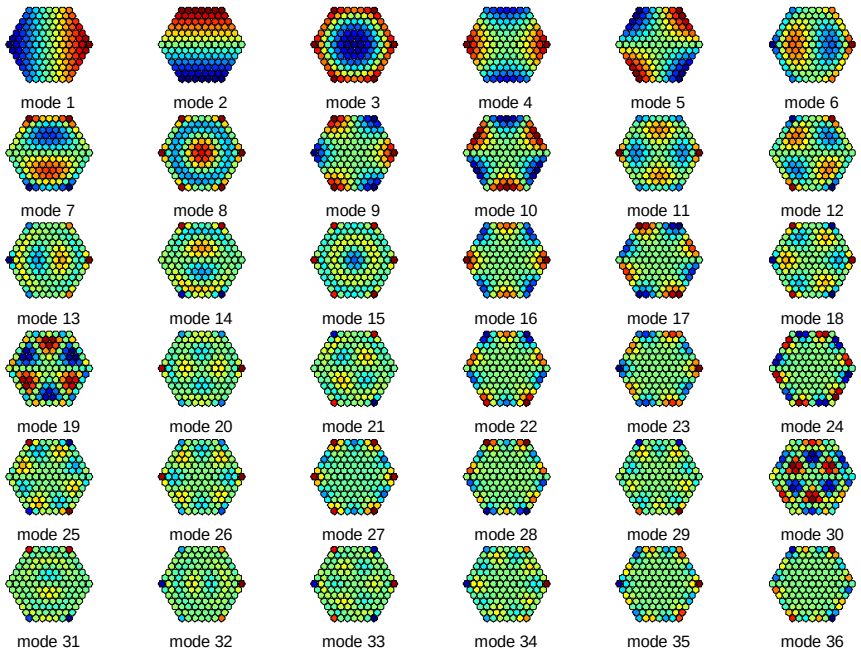


Figure 5 Visualization of Zernike basis reconstructed from Zernike derivative vectors

Adaptive Filter

Adaptive filters are a class of adaptive control laws that have been used over the past several decades in the field of active noise and vibration control. Early application of adaptive filters to the field of adaptive optics and atmospheric turbulence compensation was proposed by Ellerbroek and Rhoadarmer [1] and adaptive filter has been actively investigated by Gibson, et al. [3], [4], [8] In this paper, a SISO filtered-x (FX) Recursive Least Square (RLS) algorithm with a Finite Impulse Response (FIR) transversal filter was applied in the control space decoupled by the bases described in the previous sections and the orthogonal Zernike derivative basis described in the later section.

A block diagram of the adaptive filter, which augments an existing Proportional-Integral (PI) control feedback loop, is shown in Figure 6.

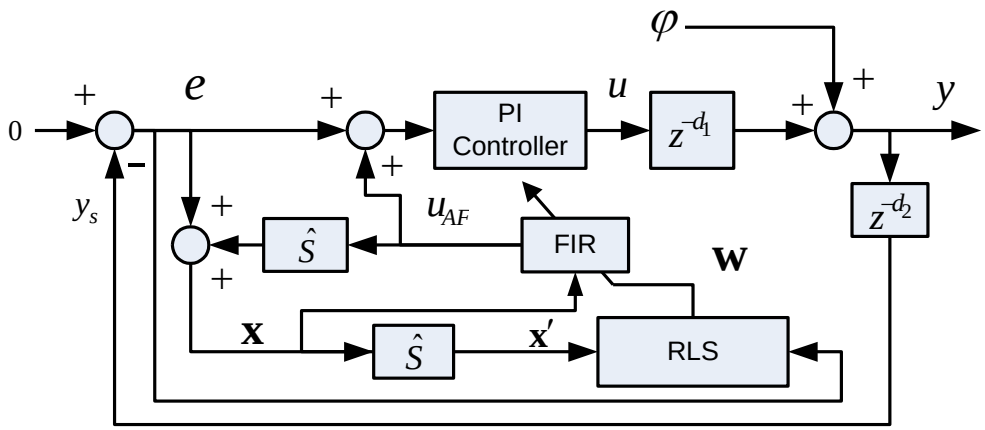


Figure 6 Block diagram of normalized Filtered-X RLS adaptive filter augmenting an existing integral control feedback loop for each uncoupled channel

The actuator and sensor delays are represented by pure time delay d_1 and d_2. The PI controller addresses the constant disturbance and the low frequency broadband disturbance while the adaptive filter addresses the disturbances which are outside the bandwidth of the PI controller. The control law of the adaptive filter loop is a FIR filter given by Eq. (10), whose input signal and adaptive coefficients $w_i(k)$ called weights are expressed as vectors defined in Eq. (11) and Eq. (12).

$$u_{AF}(k) = \sum_{i=0}^{L} w_i(k)x(k-i) = \mathbf{w}^T[k]\mathbf{x}[k] \tag{10}$$

$$\mathbf{x}(k) = \left[x(k), \quad \cdots, \quad x(k-L)\right]^T \tag{11}$$

$$\mathbf{w}(k) = \left[w_0(k), \quad \cdots, \quad w_L(k)\right]^T \tag{12}$$

The structure in Figure 6 is called feedback adaptive filter, where the input of the FIR filter called disturbance correlated signal denoted by x is estimated from the error and the adaptive filter output as:

$$x(z) = e(z) + \hat{S}(z)u_{AF}(z) = e(z) + \frac{H(z)z^{-(d_1+d_2)}}{1+H(z)z^{-(d_1+d_2)}}u_{AF}(z). \quad (13)$$

Here, $\hat{S}(z)$ is the transfer function from $u_{AF}(z)$ to $y_s(z)$ representing so-called secondary path dynamics, and $H(z)$ is the transfer function of the PI controller defined as follows with the integrator gain K_i and the proportional gain K_p:

$$H(z) = \frac{(K_i + K_p)z - K_p}{z-1} \quad (14)$$

By Eq. (13), the effect of the adaptive filter output is removed from the error, leaving the estimate of the residual error of the existing PI feedback loop with the delay. This filtered error is correlated with the disturbance to be cancelled.

Filtered-x is a technique to include the effect of the secondary path dynamics to improve stability of the adaptive filter. The reference signal goes through the secondary plant before it is fed to the adaptive algorithm as follows:

$$\mathbf{x}'(z) = \hat{S}(z)\mathbf{x}(z) = \frac{H(z)z^{-(d_1+d_2)}}{1+H(z)z^{-(d_1+d_2)}}\mathbf{x}(z). \quad (15)$$

The filter output u_{AF} is added to the input of the integral controller and the filter adaptively modifies the frequency spectrum of x such that u_{AF} cancels the disturbance observed in the error.

In the RLS algorithm, the weight vector is updated such that the following cost function is minimized:

$$\xi(n) = \sum_{i=1}^{n} \lambda^{n-i} e^2(i) \quad (16)$$

Here, λ is called the forgetting factor which places more importance on the recent data and "forgets" the data from the distant past. A value of $\lambda = 1$ implies all previous error history is included in the cost function, while smaller values excludes more past errors. A design guideline suggested for the value of λ is given as follows [9], where L+1 is the number of weights:

$$1 - \frac{1}{2L} < \lambda < 1 \quad (17)$$

Including the past history of the error usually provides faster convergence and smaller steady-state error compared with the Least Mean Square algorithm at the cost of more computational expense [10].

The update law for the weights is given by the following difference equation:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{K}(k)e(k), \quad (18)$$

where the gain vector $\mathbf{K}(k)$ is updated by:

$$\mathbf{K}(k) = \frac{\lambda^{-1}\mathbf{Q}(k-1)\mathbf{x}(k)}{[\lambda^{-1}\mathbf{Q}(k-1)\mathbf{x}(k)]^{T}[\lambda^{-1}\mathbf{Q}(k-1)\mathbf{x}(k)]+1} \tag{19}$$

The matrix $\mathbf{Q}(k) = \mathbf{R}^{-1}(k)$ is the inverse of the correlation matrix $\mathbf{R}(k)$ of the reference input calculated recursively by:

$$\mathbf{Q}(k) = \lambda^{-1}\mathbf{Q}(k-1) + \mathbf{K}(k)[\lambda^{-1}\mathbf{Q}(k-1)\mathbf{x}(k)]^{T} . \tag{20}$$

The initial condition of $\mathbf{Q}$ is a diagonal matrix whose component is determined by the expected variance of the measurement noise. For a uniform variance σ_m^2, the initial condition is given by:

$$\mathbf{Q}(0) = \frac{1}{\sigma_m^2}\mathbf{I} \tag{21}$$

III. CONTROLLER DESIGN FOR AO WITH HARDWARE LIMITATION

Modified Projection

For a static disturbance, the equation for the actuator command that eliminates the disturbance is written as follows:

$$\begin{aligned} e &= -y = -(\varphi + \mathbf{y}) = -(\varphi + \boldsymbol{\Gamma}\mathbf{u}) = \mathbf{0} \\ \varphi &= -\boldsymbol{\Gamma}\mathbf{u} \end{aligned} \tag{22}$$

For a full rank $\boldsymbol{\Gamma} \in \Re^{m\times n}$ with $n < m$, the least square solution can be obtained by multiplying both sides by $-\boldsymbol{\Gamma}^{\dagger} = -(\boldsymbol{\Gamma}\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^{T}$ as follows:

$$-\boldsymbol{\Gamma}^{\dagger}\varphi = \boldsymbol{\Gamma}^{\dagger}\boldsymbol{\Gamma}\mathbf{u} = \mathbf{u} \tag{23}$$

The residual error for this command is given as:

$$\mathbf{e}_0 = -(\varphi + \boldsymbol{\Gamma}\mathbf{u}) = -(\varphi - \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\dagger}\varphi) = -(I - \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\dagger})\varphi \tag{24}$$

Since $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\dagger} \neq \mathbf{I}$, this residual error is not zero unless $\varphi = 0$, or φ is in the null space of the matrix $(I - \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\dagger})$, which illustrates the fact that the correction of aberration introduced by a disturbance is limited by the property of the poke matrix. If the DM cannot generate the necessary mirror surface to compensate the given aberration at all points where the phase is measured, there will be residual error given by Eq. (24).

In the control space, on the other hand, the exact solution can be obtained to eliminate the error by solving the following equation for $\mathbf{u}_c$:

$$\mathbf{e}_c = -\mathbf{F}(\varphi + \boldsymbol{\Gamma}\mathbf{G}\mathbf{u}_c) = 0 \tag{25}$$

The command in control space can be obtained as:

$$-\mathbf{F}\varphi = \mathbf{F\Gamma G u}_c = \mathbf{F\Gamma}(\mathbf{F\Gamma})^{\dagger}\mathbf{u}_c = \mathbf{u}_c . \tag{26}$$

This solution $\mathbf{u}_c = -\mathbf{F}\varphi$ produces zero error in the control space but some residual error remains in the sensor space as follows:

$$\mathbf{e}_1 = -\left(\varphi + \mathbf{\Gamma}(\mathbf{F\Gamma})^{\dagger}\mathbf{u}_c\right) = -\varphi - \mathbf{\Gamma}(\mathbf{F\Gamma})\ \mathbf{F}\varphi = -\left(\mathbf{I} - \mathbf{\Gamma}(\mathbf{F\Gamma})^{?}\mathbf{F}\right)\varphi \tag{27}$$

Since $\left(\mathbf{I} - \mathbf{\Gamma}(\mathbf{F\Gamma})^{\dagger}\mathbf{F}\right) \neq \left(\mathbf{I} - \mathbf{\Gamma\Gamma}^{\dagger}\right)$ in general, the residual error $\mathbf{e}_1$ is different from the least square error $\mathbf{e}_0$. But $\mathbf{e}_0$ is the minimum error in the least square sense, and therefore $\mathbf{e}_1$ has to be larger than $\mathbf{e}_0$, i.e., $|\mathbf{e}_1|^2 > |\mathbf{e}_0|^2$.

This inconvenience, however, can be fixed by replacing $\mathbf{F}$ with $\mathbf{F}' = \mathbf{F\Gamma\Gamma}^{\dagger}$ in equation (26) while keeping $\mathbf{G} = (\mathbf{F\Gamma})^{\dagger}$. The $\mathbf{\Gamma\Gamma}^{\dagger}$ term does not affect the decoupling of the open-loop path as shown in the following:

$$-\mathbf{F}'\varphi = \mathbf{F}'\mathbf{\Gamma G u} = \mathbf{F\Gamma}(\mathbf{\Gamma}^{\dagger}\mathbf{\Gamma})(\mathbf{F\Gamma})^{\dagger}\mathbf{u}'_c = \mathbf{F\Gamma}(\mathbf{F\Gamma})^{\dagger}\mathbf{u}'_c = \mathbf{u}'_c \tag{28}$$

But the solution in the control space is now changed to $\mathbf{u}'_c = -\mathbf{F\Gamma\Gamma}^{\dagger}\varphi$.

Applying this solution to Eq. (27) produces the least square error in the sensor space as shown below, provided that $\mathbf{F\Gamma}$ is square and invertible:

$$\begin{aligned}\mathbf{e}'_2 &= -\left(\varphi + \mathbf{\Gamma}(\mathbf{F\Gamma})^{\dagger}\mathbf{u}'_c\right) = -\varphi + \mathbf{\Gamma}(\mathbf{F\Gamma})\ \mathbf{F\Gamma\Gamma}^{\dagger}\varphi = -\varphi + \mathbf{\Gamma}(\mathbf{F\Gamma})\ (\mathbf{F\Gamma})\mathbf{\Gamma}^{?}\varphi \\ &= -\left(\mathbf{I} - \mathbf{\Gamma\Gamma}^{\dagger}\right)\varphi = \mathbf{e}_0\end{aligned} \tag{29}$$

From a physical point of view, the term $\mathbf{\Gamma\Gamma}^{\dagger}$ removes the components of the disturbance that are orthogonal to the poke matrix column vector space, which is the subspace the deformable mirror does not have any influence.

Orthogonalization of Zernike derivative basis

The Zernike polynomial is orthogonal and normal over the interior of a unit circle only for a continuous case. When it is evaluated at discrete points in the unit circle, the resulting vectors are not orthogonal. In addition, the derivative vectors (8) is not orthogonal even for a continuous case. Since an orthogonal basis is desired for modal control, a simple technique to obtain orthogonal Zernike derivative vectors which preserve the spatial characteristics of the low order modes to a certain extent can be investigated. The objective here is to obtain a basis that resembles Zernike and not to determine the true Zernike coefficients.

Let *dZ* be a matrix of Zernike derivative vectors which is sorted in ascending order from the left to right. Construct an arbitrary diagonal weight matrix whose

diagonal components are ordered in an ascending manner. An example of weight is given as follows:

$$w_k = \exp(a(N-k+1))/\exp(N) \tag{30}$$

where N is the number of Zernike basis and a is some positive constant. Post-multiplying the weight matrix to the Zernike derivative matrix and taking the singular value decomposition give the following three matrices:

$$\mathbf{U}_z \mathbf{\Sigma}_z \mathbf{V}_z^T = \mathrm{SVD}\{d\mathbf{Z}\mathbf{W}\} \tag{31}$$

The column vectors of $\mathbf{U}_z$ is orthogonal and the shapes of the lower order modes are similar to those of the original Zernike derivative vectors. The column vectors of $\mathbf{U}_z$ is referred to hereafter as the orthogonal Zernike derivative basis $d\tilde{\mathbf{Z}} = \mathbf{U}_z$. This method can also be applied to obtain an orthogonal Zernike basis by replacing $d\mathbf{Z}$ in Eq. (31) with the discretized Zernike basis matrix $\mathbf{Z}$.

Figure 7 shows the visualization of the orthogonality of the obtained basis vectors for the weights shown in Figure 8. The plot on the left is the components of $d\tilde{\mathbf{Z}}^T d\tilde{\mathbf{Z}}$ and the plot on the right shows those of $d\mathbf{Z}^T d\mathbf{Z}$. Each basis vectoris normalized to have Euclidean norm 1. The matrix $d\tilde{\mathbf{Z}}^T d\tilde{\mathbf{Z}}$ is an identity matrix indicating that the column vectors of $d\tilde{\mathbf{Z}}$ are mutually orthogonal. The product $d\mathbf{Z}^T d\mathbf{Z}$, on the other hand, has non-zero off-diagonal components indicating that the column vectors of $d\mathbf{Z}$ are not exactly orthogonal.

If $d\tilde{\mathbf{Z}}^T \mathbf{\Gamma}$ is full rank, setting $\mathbf{F} = d\tilde{\mathbf{Z}}^T$ and $\mathbf{G} = \left(d\tilde{\mathbf{Z}}^T \mathbf{\Gamma}\right)^{\dagger}$ will decouple the open loop path observed from the control space. The coefficient of the original Zernike basis can be obtained from the coefficient of the orthogonal Zernike derivative basis $\tilde{\mathbf{a}} = d\tilde{\mathbf{Z}}^T \mathbf{e}_s$ as follows:

$$\mathbf{a} = d\mathbf{Z}^{\dagger}\mathbf{e}_s = \mathbf{W}\mathbf{V}_z\mathbf{\Sigma}_z^{\dagger} \quad \tilde{\mathbf{a}} \tag{32}$$

Figure 9 is the visualization of the corresponding wavefront shapes of the first 36 Zernike derivative basis vectors excluding piston mode. It can be seen that the spatial frequency characteristics are similar to the Zernike derivative shown in Figure 5 for lower order modes.

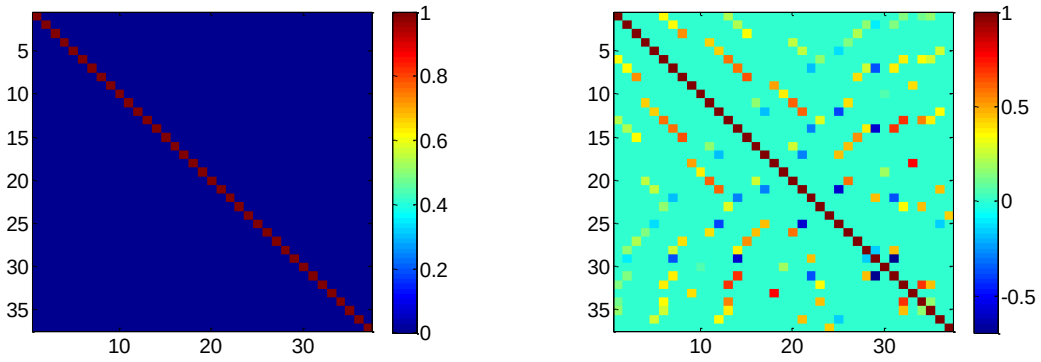


Figure 7 Orthogonality of the orthogonal Zernike and original Zernike derivative basis

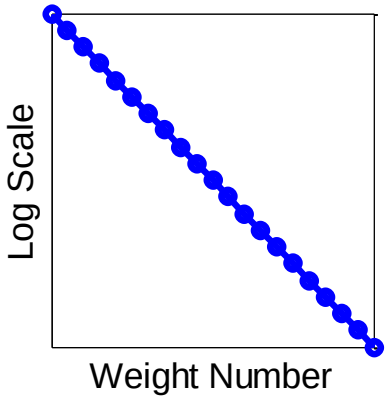


Figure 8 Example of diagonal component of weight matrix

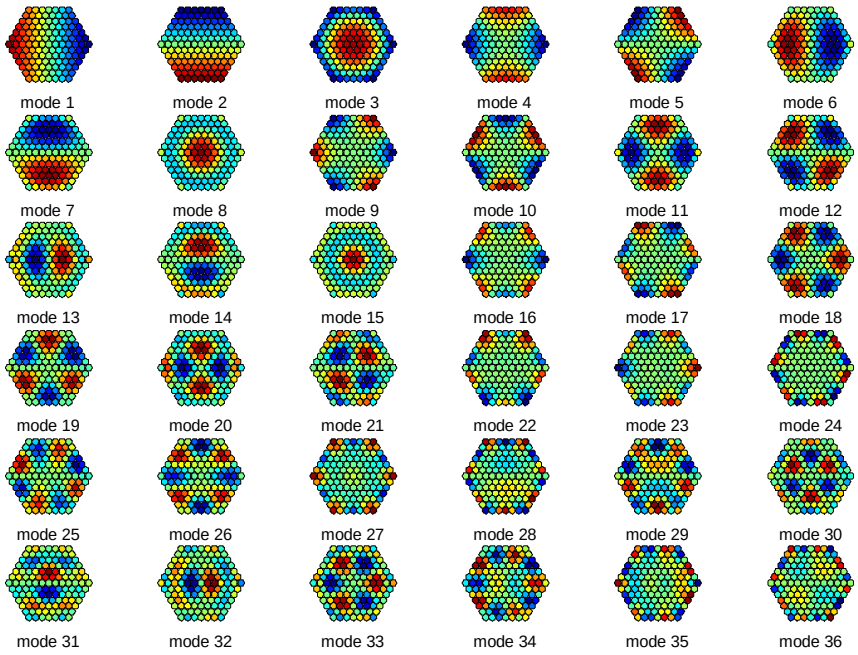


Figure 9 Visualization of the phase reconstructed from the orthogonal Zernike derivative basis

Anti-windup control

When saturation of the DM occurs due to a temporarily surge of the disturbance, the state of the integrator in the PI controller can grow, or "wind up", and the controller output continues to saturate the actuator even after the disturbance surge is gone. This so-called windup problem is inherent in a system with integral controller, and a technique called anti-windup control has been developed in control engineering. Figure 10 shows the block diagram of a simple anti-windup control applied to the control system considered in this paper.

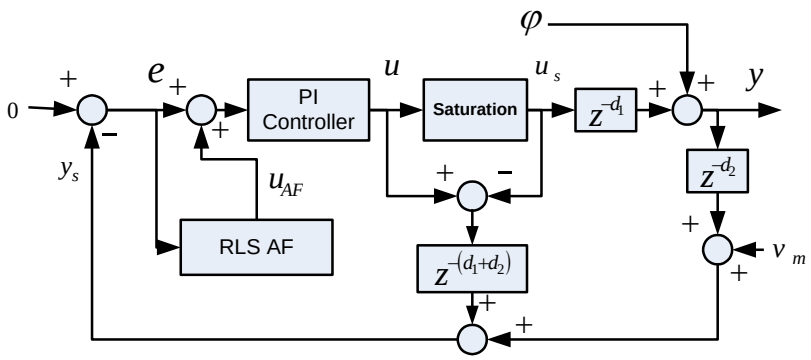


Figure 10 PI and adaptive filter controller with anti-windup feedback signal

The anti-windup control estimates the effect of the controller suppressed by the actuator saturation and compensates the error fed to the controller to avoid the overcorrection by the controller. It cannot prevent actuator saturation, but the effect of the saturation is now limited to the actual duration of the time when a large disturbance is present to cause actuator saturation. Once the disturbance surge retreats, the controller performance goes back the level where no saturation is involved.

IV. EXPERIMENTS AND RESULTS

Adaptive Optics Testbed

Figure 11 shows a picture of the AO testbed used for the experiment. It consists of a laser unit, a SH WFS, a DM, and two liquid crystal (LC) spatial light modulators (SLMs) to project a dynamic phase aberration. One science camera is placed to capture the beam image. It is a CCD camera with a resolution of 640 x 480 pixels and an 8-bit frame rate of 75 fps. The laser used in the experiment is a continuous wave Helium Neon Class IIIa laser with output power of 0.5 mW, operating at a wavelength of 633 nm.

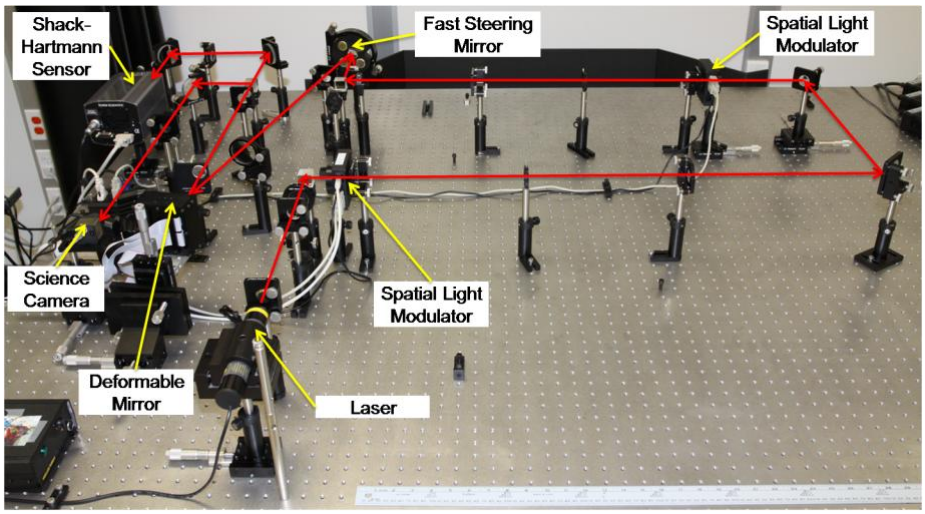


Figure 11 Schematic of laboratory system

Wavefront Sensor

The SH WFS consists of 127 lenslets in a hexagonal pattern and a CMOS monochrome camera with a resolution of 640 x 480 pixels and an 8-bit frame rate of 60 fps. In order to obtain the actual slopes, the measured offset of the spot centers have to be divided by the focal length of the lenslets. In this study, the slope error measured by the SH WFS is evaluated only in relative terms and the spot center deviation in the unit of pixel is used as the slope error.

It is common to remove so-called tip/tilt modes by a fast steering mirror instead of the deformable mirror as the tip and tilt error is often much larger than the deformable mirror stroke. In the experiments, the tip/tilt modes are ignored by subtracting the mean of the *x* and *y* slopes from the error instead of forming a separate tip/tilt correction loop.

Deformable Mirror

The DM is a 37-channel micro-machined membrane deformable mirror (MMDM) and controlled by applying an array of voltages to electrodes on the back surface of the mirror. The membrane mirror is fixed on the outside rim and the applied voltage moves the mirror surface in one direction from a flat reference producing concave shapes. In order to allow bidirectional control, the mirror is initially set at a biased position in the middle of its range of deflection. The input signal to the DM normalized from -1 to 1 is converted so that the quadratic relationship between the voltage and the mirror surface movement is compensated and the zero command corresponds to the bias position. The DM saturates for input outside of -1 to 1 range.

The poke matrix used in the control was measured by applying the input command to a single channel and observing the output of the WFS. For each channel, poke commands with various magnitudes were applied to cover the entire input range and a least square method was applied to obtain the mean value of the poke matrix. Because the DM has the fixed rim, the DM cannot produce the piston mode which cannot be observed by SH WFS. As a result, the poke matrix has the full rank.

SLM and Disturbance Generation

Only one of the SLM located in the system pupil plane conjugate to the wavefront sensor, the deformable mirror is used, and no aberration is applied to the other SLM. The resolution of the SLM is 800x600 pixels and the maximum operational rate is 33Hz. It consists of a diffraction grating that modulates the incoming wavefront by π radians. To increase the modulation range to a full 2π, a Fourier filter in the form of an iris or aperture stop is placed in the beam to select either the +1 or -1 diffractive order. Alignment biases are applied to separate the diffractive orders enough to pass through the beam of the desired order.

The SLM is driven by software developed by the Naval Research Laboratory (NRL) [11]. The algorithm generates the phase aberration by augmented K-L polynomial expansion shown in Eq. (33), which includes Tatarskii's assumption of a Gaussian random distribution in phase variances due to turbulence.

$$\text{Wavefront}(\rho,\theta)=\sum_{i=1}^{M}(1+X_i(t))a_iK_i(\rho,\theta) \tag{33}$$

Here, M is the number of K-L modes and a_i are the coefficients representing the weights given to each mode. The coefficients are calculated based on the specified parameters such as the telescope diameter, D, and the atmospheric coherence length or Fried parameter, r_0, in addition to the Zernike-Kolmogorov residual errors measured experimentally by Fried [12] and calculated by Noll [13]. The Karhunen-Loève modes are used because they contain a statistically independent set of coefficients based on Zernike modes, and they are often used in turbulence simulation.

The X_i is a continuous function of time generated by interpolating the Gaussian distribution random numbers with spline curve to obtain smooth transition between frames.

The disturbance used in the experiments is generated based on the atmospheric profile of a telescope aperture of 1 m diameter and an atmospheric coherence length of $r_0=15$ cm, representing an atmosphere of medium strength. The actual aberration observed by the WFS is slightly scaled down due to the optical alignments of the testbed. The Zernike modes contained in the aberration are 18 modes and the interpolation points for X_i are 7 or 8 intervals between the random numbers. The random number generation and spline curve fitting are repeated for each mode and the phase screen that combines all modes are applied on the SLM operating at the sample rate of 33Hz.

Controller

A desktop computer is used to control the deformable mirror, the SH WFS, and the science camera, and another computer is used to control the SLMs to produce the disturbance. The control algorithm is implemented in Matlab script and executed under a Microsoft Windows XP operating system. Although this system does not provide real-time control of the hardware, an external driver program to obtain the images from the WFS camera sets the upper bound of the sample rate of the control loop and the sample rate is maintained in a usable range.

Modified Modal Decomposition

Figure 12 shows the RMS of the sensor space error when the conventional modal decomposition using **F** and **G** described in Section II was applied. The aberration is static and the controller is a PI controller, and all 37 modes are controlled except for the orthogonal Zernike derivative and Zernike derivative bases with which tip and tilt modes are excluded. While the poke matrix basis and SVD basis converge to a steady state value, the orthogonal Zernike derivative basis and regular Zernike derivative basis are having difficulty to settle to a constant value. Figure 13 shows the results of the same experiments when the modified projection is applied. The errors by all bases now converge to the same value which is supposed to be the minimum error that can be achieved with the given physical system.

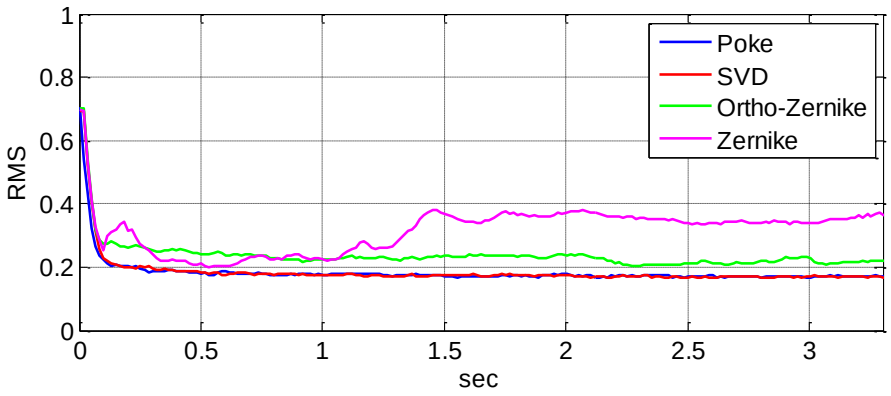


Figure 12 RMS of the error vector components no filtering of the error is applied

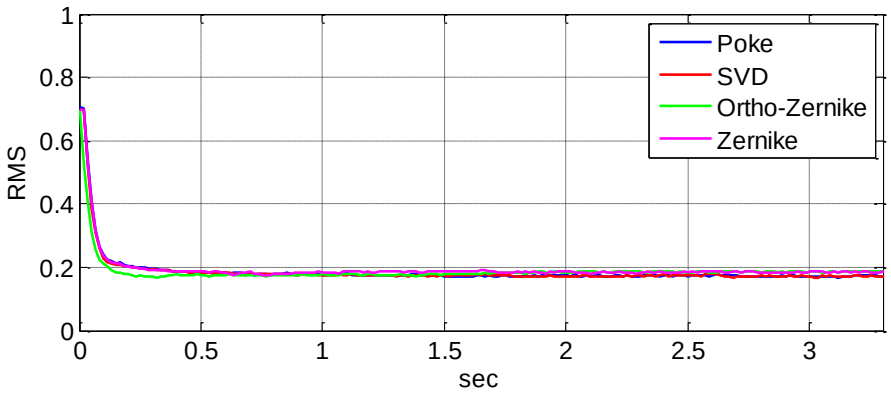
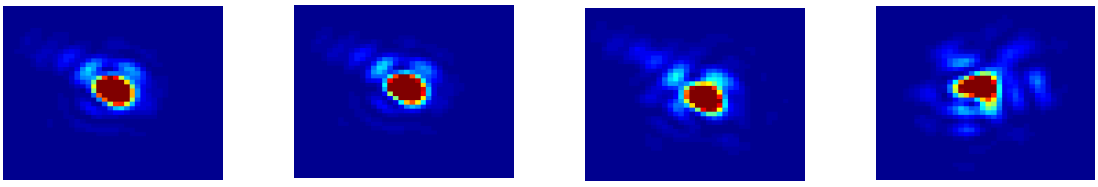


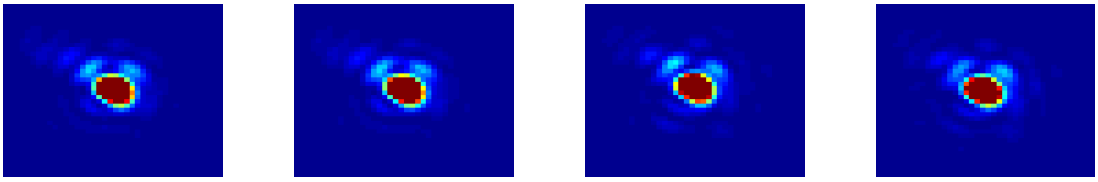
Figure 13 RMS of the error vector components when the filtering of the error is applied

Figure 14 and Figure 15 show the corresponding science camera images at steady state.



Poke Basis **SVD Basis** **Orthogonal Zernike Derivative** **Zernike Derivative**

Figure 14 Image of the beam with conventional projection of the error



Poke Basis **SVD Basis** **Orthogonal Zernike Derivative** **Zernike Derivative**

Figure 15 Image of the beam with modified projection method

Significant improvement by the modified projection method can be observed for the Zernike type basis.

Anti-Windup Controller

The effectiveness of the anti-windup control was evaluated with a dynamic phase aberration whose RMS value observed by the WFS is shown in Figure 16. The performance of the PI and the PI with RLS adaptive filter control were obtained for the four different bases. All 37 modes are controlled by the poke and SVD basis, and 35 modes without the piston, tip, and tilt modes are controlled by the orthogonal Zernike derivative and Zernike derivative bases. The modified projection is applied in all cases. Table 1 shows the parameters of the controllers which were obtained from preliminary trials.

Table 1 Controller parameters for dynamic aberration control

Full (37) modes	Poke Basis	SVD Basis	Ortho-Zernike	Zernike
PI controller	$K_i = 0.2$ $K_p = 0.05$	$K_i = 0.2$ $K_p = 0.05$	$K_i = 0.2$ $K_p = 0.05$	$K_i = 0.2$ $K_p = 0.05$
RLS AF	$\lambda = 0.9999$ $\sigma_m^2 = 4$ $L = 15$	$\lambda = 0.9999$ $\sigma_m^2 = 5$ $L = 15$	$\lambda = 0.9999$ $\sigma_m^2 = 300$ $L = 15$	$\lambda = 0.9999$ $\sigma_m^2 = 150$ $L = 15$

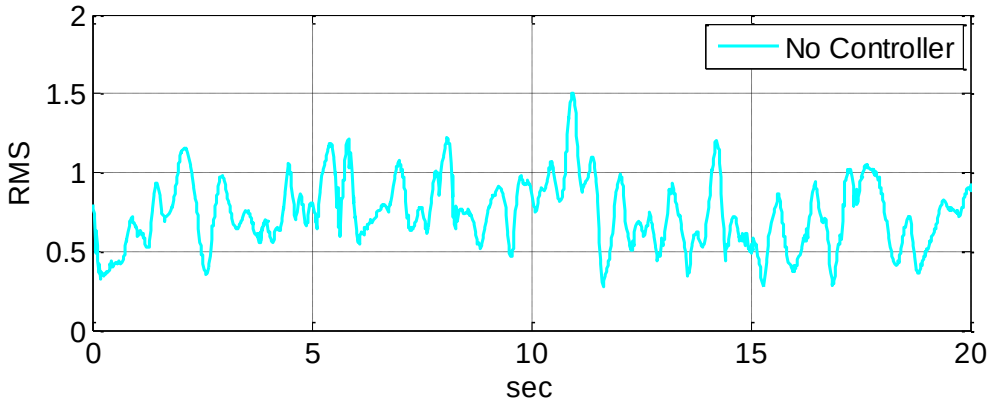


Figure 16 RMS of the sensor space error vector without applying a feedback controller

Figure 17 through Figure 1920 show the sensor space error RMS by the PI controller with and without the anti-windup control for the poke, SVD, Zernike derivative and orthogonal Zernike derivative, respectively. Figure 23 through Figure 21 show the result of PI with RLS adaptive filter control.

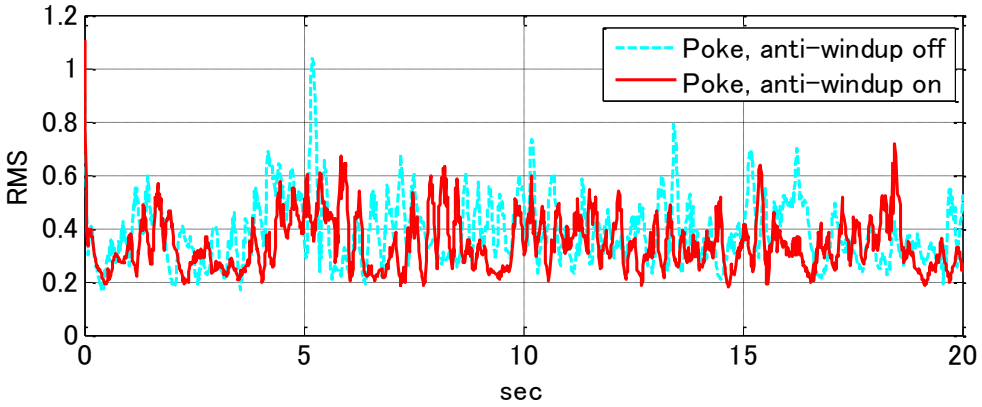


Figure 17 RMS of the sensor space error vector by the PI controller with poke basis

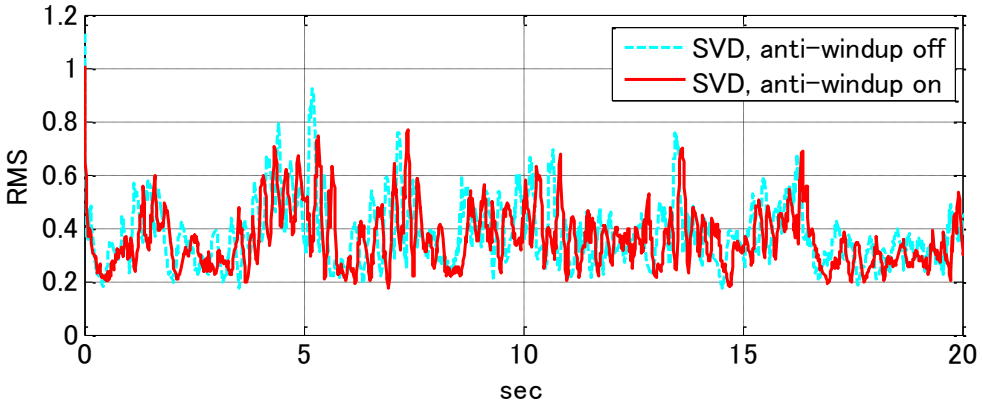


Figure 18 RMS of the sensor space error vector by the PI controller with SVD basis

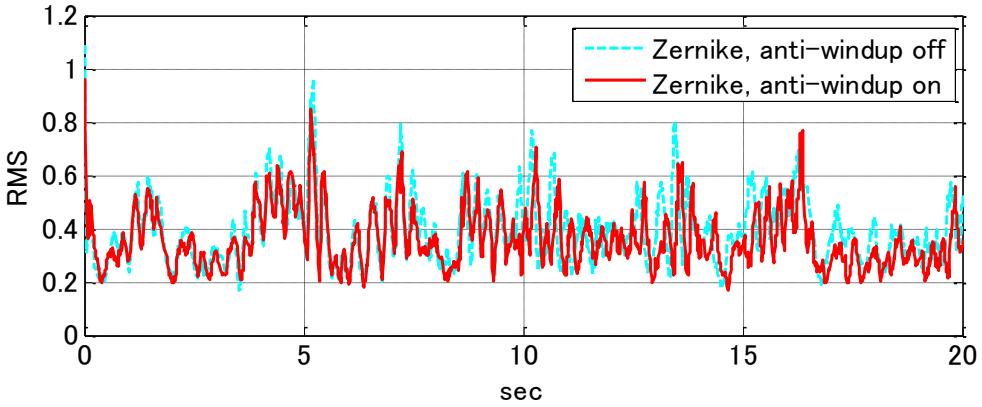


Figure 19 RMS of the sensor space error vector by the PI controller with Zernike basis

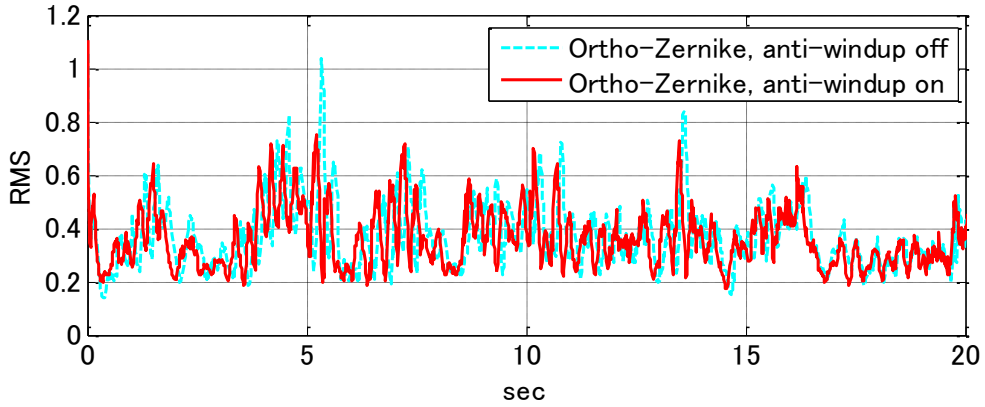


Figure 20 RMS of the sensor space error vector by the PI controller with orthogonal Zernike derivative basis

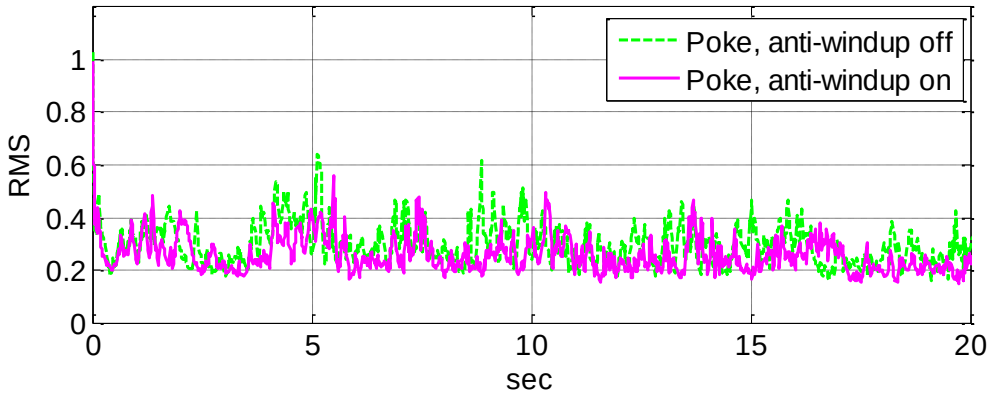


Figure 21 RMS of the sensor space error vector by the PI and RLS adaptive filter controller with poke basis

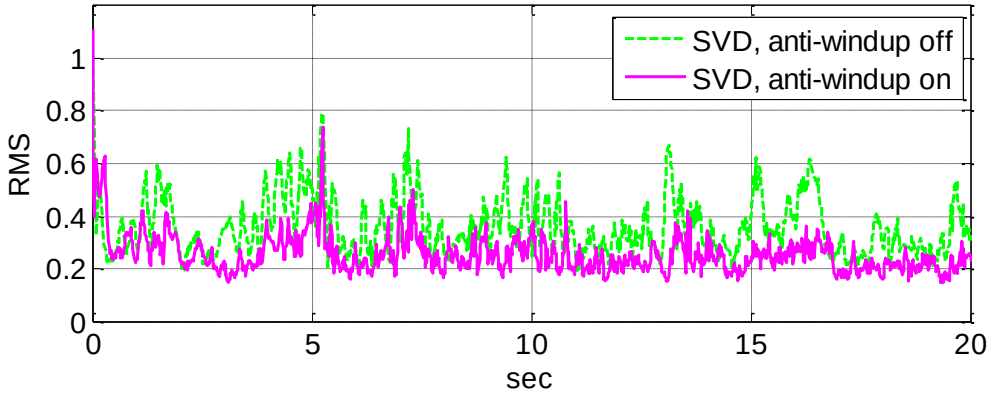


Figure 22 RMS of the sensor space error vector by the PI and RLS adaptive filter controller with SVD basis

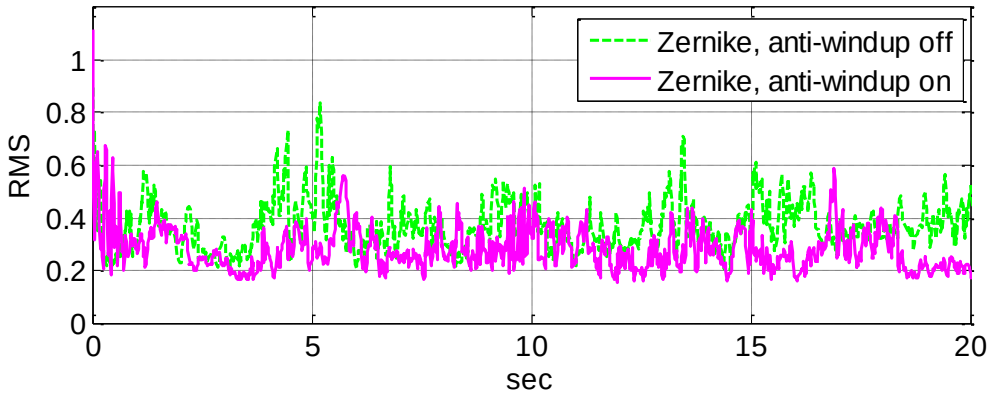


Figure 23 RMS of the sensor space error vector by the PI and RLS adaptive filter controller with Zernike basis

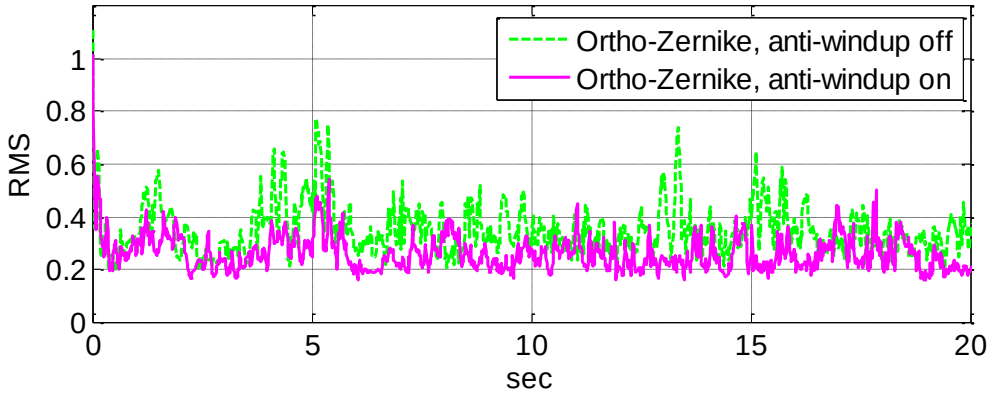


Figure 24 RMS of the sensor space error vector by the PI and RLS adaptive filter controller with orthogonal Zernike derivative basis

Figure 25 and Figure 26 show the time average of the RMS of the errors shown in the previous figures. In both PI and PI with RLS adaptive filter cases, the error is reduced by the anti-windup control.

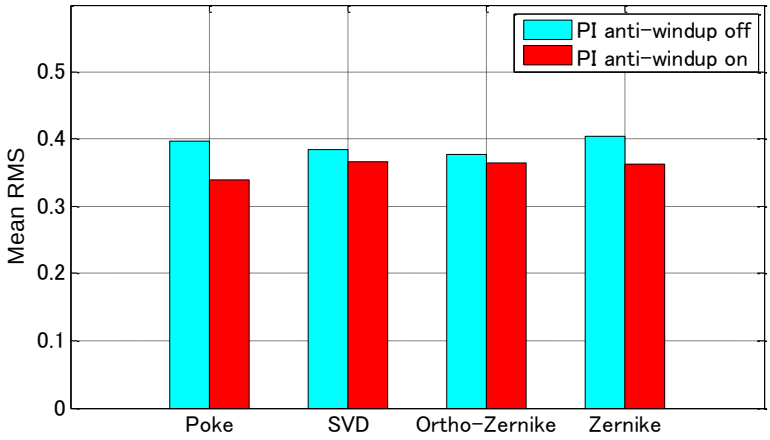


Figure 25 Summary of error reduction by anti-windup with PI controller

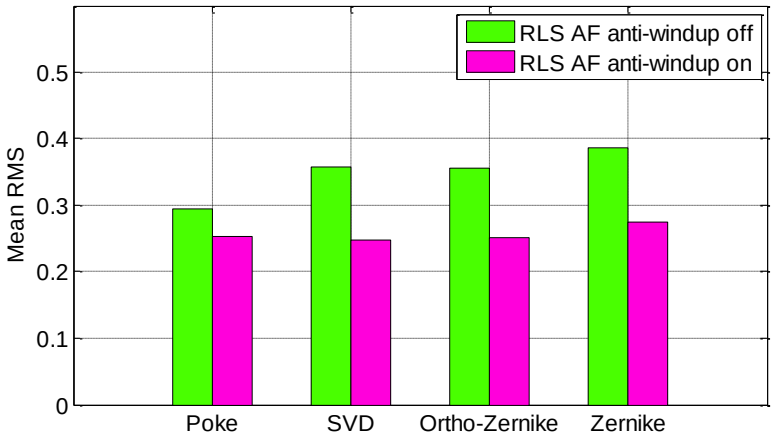


Figure 26 Summary of error reduction by anti-windup with PI and RLS adaptive filter

More significant improvement is obtained when the RLS adaptive filter is used, and it is because the adaptive filter attempts to attenuate the disturbance at a wider frequency range than the PI controller which leads to more frequent actuator saturation.

Analysis of the effect of mode reduction

Figure 27 and Figure 28 show the time average RMS of the sensor space error vector plotted against the number of basis vectors used for the PI controller and PI with RLS adaptive filter, respectively. The first three modes, namely piston, tip, and tilt, are removed from the Orthogonal Zernike derivative basis and Zernike derivative basis control and 35 is the maximum number of modes addressed for these bases. The modified projection and anti-windup control are both applied.

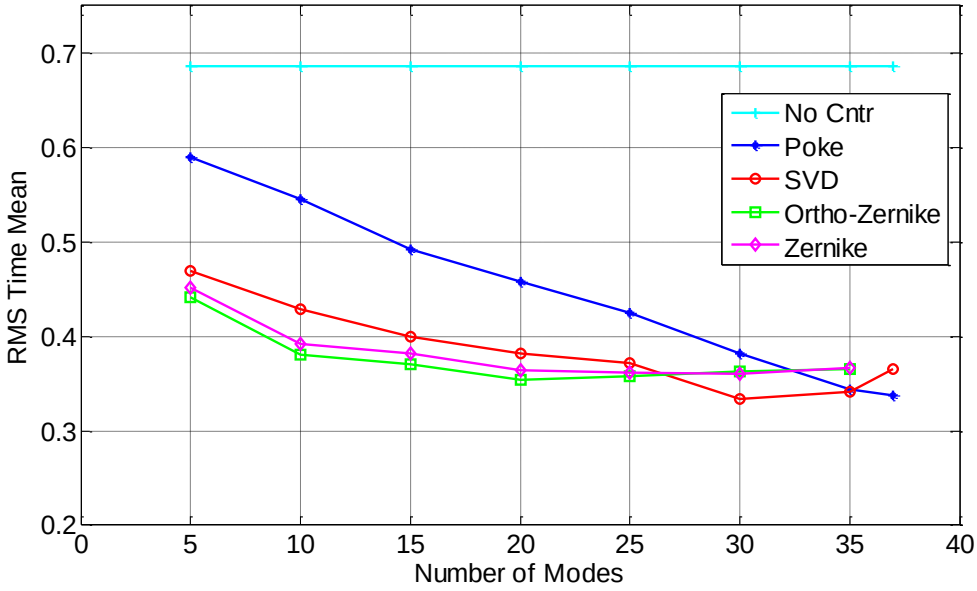


Figure 27 Time mean of the sensor space error RMS by PI controller

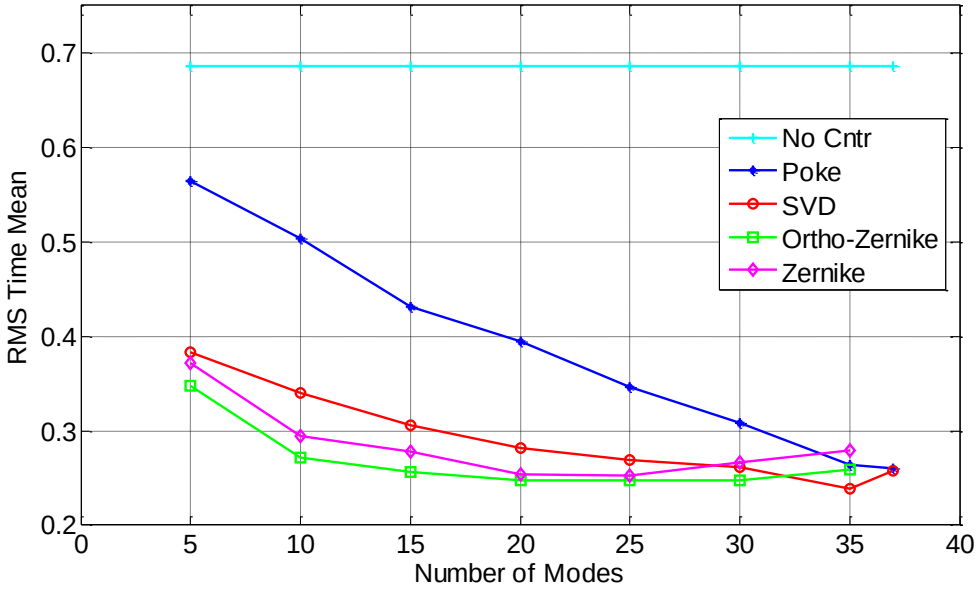


Figure 28 Time mean of the sensor space error RMS by PI and RLS adaptive filter controller

For the poke matrix basis, the error increases as the number of addressed modes reduces, while the curve of the SVD basis is less steep. For the orthogonal Zernike derivative and Zernike derivative bases, the error actually reduces as the addressed modes are reduced down to 20 modes before the error increases with the reduction of the modes. This result is further analyzed by investigating the decomposition of the disturbance by each basis.

Figure 29 shows the magnitude of the error projected on each basis. The value is normalized by the largest magnitude for each mode. It can be seen that the error is concentrated in the lower order modes for the SVD and Zernike type modes, but no such concentration is observed for the poke matrix basis. In poke basis, the higher modes are also significance in the phase error and reduction of those modes directly impacts the resulting error.

For SVD, the basis vectors are ordered in terms of the "gain" of each mode, and ignoring the higher order modes, which is harder for the DM to control even in the full-mode case. Also, the projection of the disturbance on the SVD basis is more concentrated in the lower modes, and the error increase by mode reduction is much smaller than that of the poke matrix.

In this experiment, the disturbance is generated by the Karhunen-Loève expansion which is based on the Zernike polynomials up to 18th order, and the coefficients are defined in the disturbance generation to reflect the atmospheric aberration's spatial frequency spectrum. As a result, the concentration of the disturbance to the lower order modes is even stronger in the Zernike type bases than the SVD basis case and these bases can produce less error with less modes addressed. It was found the orthogonal Zernike derivative basis performs slightly better than the normal Zernike basis with the PI with RLS adaptive filter control.

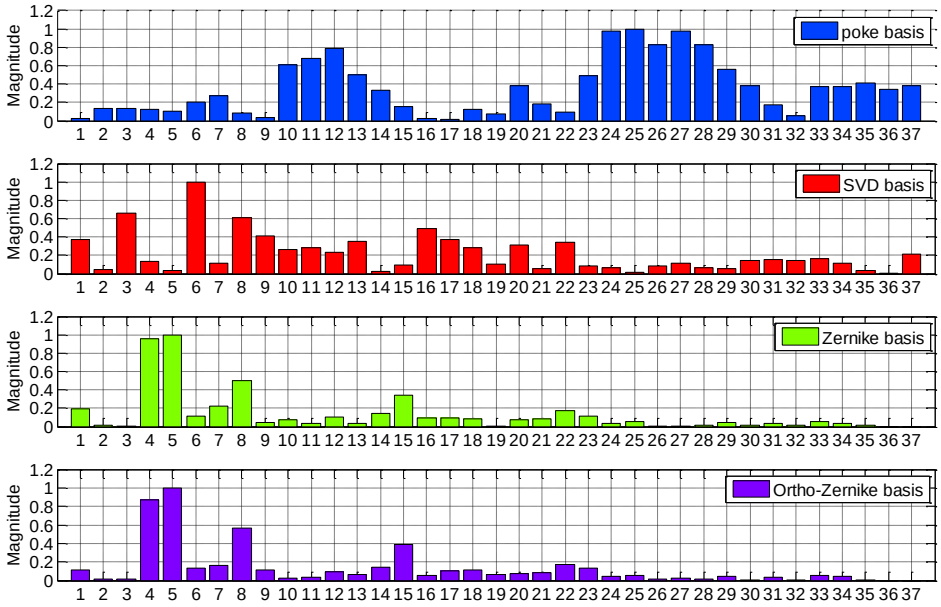


Figure 29 Normalized magnitude of the error projected on each basis

V. CONCLUSION

In this paper, practical control techniques for modal control of AO systems focuses on the issues arising from the limitation of the hardware, namely, the spatial resolution and finite stroke of the DM are proposed and the effectiveness has been experimentally evaluated. An orthogonal basis approximating the Zernike derivative basis is also proposed and the performance has been evaluated in comparison with common bases found in AO control application, namely, the poke basis (pseudo inverse method), the SVD basis, and the discretized Zernike derivative basis.

For a deformable mirror with less actuators than the sensor measurement points, there is a subspace of the sensor vector space where the actuator simply cannot have any influence, and the effort by the controller to address the error components in this subspace can lead to a performance degradation. In the proposed modified projection method, the error is first projected to the subspace that can be controlled by the actuator by the pseudo inverse matrix, and then projected back to the sensor space by applying the poke matrix. The procedure does not affect the decoupling of the control path but the error components in the uncontrollable subspace are effectively filtered out from the error observed by the controller, provided that the poke matrix is full rank. This method would not have any effect for bases that have an explicit relationship with the poke matrix, i.e., the geometry of the DM and the WFS, such as poke and SVD bases. Experimental results show that the error by the Zernike type bases converge to a constant minimum steady state error produced by the poke and SVD bases with the proposed method, while they have difficulty settling to a constant steady state error with the conventional method.

In adaptive optics control, an integral controller is often used to eliminate the constant disturbance. In the presence of actuator saturation, however, it can cause a so-called windup problem which results in degradation of the performance. An anti-windup control technique is proposed for AO systems with actuator saturation and experimental results showed the proposed method is effective in limiting the saturation effect,

especially with an RLS adaptive filter which tends to saturate the DM more often due to a more aggressive attempt to attenuate the disturbance in a wider frequency range than the PI controller.

Orthogonality is one of the desired properties for the basis used in modal control, but the Zernike derivative basis used for control of Zernike modes is not orthogonal. In the proposed orthogonalization method, the Zernike derivative basis is orthogonalized using SVD with a diagonal weighting matrix and the resulting basis approximates the spatial characteristics of the original Zernike modes. The diagonal weighting matrix serves as a means to control the relative accuracy of each mode. For the logarithmically reducing weights, the discrepancy between the phase shapes of the orthogonalized Zernike modes and those of the original Zernike modes increases as the order goes higher. The experimental result shows that the orthogonal Zernike derivative basis combined with the modified projection and the anti-windup methods achieves better performance than other bases considered in this paper with both PI and PI with RLS adaptive controllers when the number of addressed modes is reduced.

As the application of AO technology expands, more cost effective AO systems will be demanded which requires the controller to make most out of the given hardware with limited capabilities. Therefore, the study of practical control techniques such as the ones presented in this paper will carry more importance as the research moves from demonstrating the feasibility to maturing the technology for wider applications.

ACKNOWLEDGEMENT

This research was conducted while the first author held the National Research Council Research Associateship at the Naval Postgraduate School.

REFERENCES

[1] B. L. Ellerbroek and T. A. Rhoadarmer, "Real-time adaptive optimization of wavefront reconstruction algorithms for closed-loop adaptive optical systems," in *Proceedings of SPIE*, vol. 3353, pp. 1174–1185, 1998.

[2] T. A. Rhoadarmer, L. M. Klein, J. S. Gibson, N. Chen, and Y.-T. Liu, "Adaptive control and filtering for closed-loop adaptive-optical wavefront reconstruction," *Proceedings of SPIE*, vol. 6306, p. 63060E, 2006.

[3] Y.-T. Liu and J. S. Gibson, "Adaptive control in adaptive optics for directed-energy systems," *Optical Engineering*, vol. 46, no. 4, p. 046601, 2007.

[4] S. Monirabbasi and J. S. Gibson, "Adaptive control in an adaptive optics experiment," *J. Opt. Soc. Am. A*, vol. 27, no. 11, pp. A84–A96, Nov. 2010.

[5] F. Roddier, Ed., *Adaptive Optics in Astronomy*. Cambridge University Press, 1999.

[6] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge University Press, 1999.

[7] W. H. Southwell, "Wave-front estimation from wave-front slope measurements," *Journal of the Optical Society of America*, vol. 70, no. 8, pp. 998–1006, Aug. 1980.

[8] J. S. Gibson, C.-C. Chang, and B. L. Ellerbroek, "Adaptive optics: wave-front correction by use of adaptive filtering and control," *Applied optics*, vol. 39, no. 16, pp. 2525–38, Jun. 2000.
[9] S. O. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 2001.
[10] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*. Wiley-Interscience, 1996.
[11] C. C. Wilcox, "Optical phase aberration generation using a liquid crystal spatial light modulator," Ph.D. Dissertation, University of New Mexico, Albuquerque, 2009.
[12] D. L. Fried, "Statistics of a Geometric Representation of Wavefront Distortion," *Journal of the Optical Society of America*, vol. 55, no. 11, pp. 1427–1431, 1965.
[13] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *Journal of the Optical Society of America*, vol. 66, no. 3, pp. 207–211, 1976.