

# Suppressing Anomalous Localized Waffle Behavior in Least Square Wavefront Reconstructors

*D. Gavel*

This article was submitted to  
SPIE (Society of Photo-Optical Instrumentation Engineers),  
Adaptive Optical System Technologies II,  
Waikoloa, HI  
08/22/02 – 08/28/02

*U.S. Department of Energy*

Lawrence  
Livermore  
National  
Laboratory

**October 8, 2002**

## DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

This report has been reproduced directly from the best available copy.

Available electronically at <http://www.doc.gov/bridge>

Available for a processing fee to U.S. Department of Energy  
And its contractors in paper from  
U.S. Department of Energy  
Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831-0062  
Telephone: (865) 576-8401  
Facsimile: (865) 576-5728  
E-mail: [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)

Available for the sale to the public from  
U.S. Department of Commerce  
National Technical Information Service  
5285 Port Royal Road  
Springfield, VA 22161  
Telephone: (800) 553-6847  
Facsimile: (703) 605-6900  
E-mail: [orders@ntis.fedworld.gov](mailto:orders@ntis.fedworld.gov)  
Online ordering: <http://www.ntis.gov/ordering.htm>

OR

Lawrence Livermore National Laboratory  
Technical Information Department's Digital Library  
<http://www.llnl.gov/tid/Library.html>

# Suppressing Anomalous Localized Waffle Behavior in Least Squares Wavefront Reconstructors

Donald Gavel\*

Lawrence Livermore National Laboratory

## ABSTRACT

A major difficulty with wavefront slope sensors is their insensitivity to certain phase aberration patterns, the classic example being the waffle pattern in the Fried sampling geometry. As the number of degrees of freedom in AO systems grows larger, the possibility of troublesome waffle-like behavior over localized portions of the aperture is becoming evident. Reconstructor matrices have associated with them, either explicitly or implicitly, an orthogonal mode space over which they operate, called the singular mode space. If not properly preconditioned, the reconstructor's mode set can consist almost entirely of modes that each have some localized waffle-like behavior. In this paper we analyze the behavior of least-squares reconstructors with regard to their mode spaces. We introduce a new technique that is successful in producing a mode space that segregates the waffle-like behavior into a few "high order" modes, which can then be projected out of the reconstructor matrix. This technique can be adapted so as to remove any specific modes that are undesirable in the final reconstructor (such as piston, tip, and tilt for example) as well as suppress (the more nebulously defined) localized waffle behavior.

**Keywords:** Wavefront reconstructors, waffle mode

## 1. INTRODUCTION

The Hartmann-Shack sensor provides measurements of the subaperture-averaged wavefront slope at sample positions within an aperture. In principle, given these measurements it is possible to recreate the original wavefront surface under the assumption that the wavefront is continuous across sample subaperture boundaries. Several authors have developed methods for wavefront reconstructing<sup>1,2,3</sup>. In this paper we are chiefly concerned with the least-squares approach, where the objective is to find the wavefront that minimizes a norm of differences between Hartmann sensor measurements and the measurements that would be obtained from the estimated wavefront. This approach is quite commonly used as the basis for closed loop compensation of atmospheric turbulence in adaptive optics systems<sup>4,5</sup>. The least-squares fit is a reasonable objective when the signal-to-noise ratio of the measurement is high, since there is high confidence placed in the measurements when the estimate is made to fit them as closely as possible. When the signal-to-noise is low, there are refinements to the least-squares approach, such as weighted-least-squares and regularized weighted-least-squares, which can take into account relative confidence in the measurements and/or wavefront prior information.

A major difficulty with Hartmann sensors is the existence of blind or null modes, that is, wavefront functions that can yield zero or very small response in the Hartmann sensor output. The classic example is the waffle pattern in the Fried sample geometry<sup>3</sup>. Waffle mode is essentially a repeated astigmatic pattern at the period of the Hartmann sampling over the entire aperture. This wavefront has zero average slope over every subaperture and hence produces zero Hartmann sensor output. The null mode set represents anomalous wavefront functions that will satisfy the least-squares-fit criterion exactly, but may adversely affect the far-field imaging performance of an adaptive optics system. Imaging performance is typically characterized in terms of Strehl ratio, the far-field on-axis intensity relative to that of an unaberrated beam. The Marechal approximation<sup>6</sup> states that the focal plane Strehl is monotonically related to the mean-square wavefront residual. Therefore minimizing the mean-square wavefront residual is a better objective than mean-square fit to measurements, but, because of the Hartmann sensor's null space, the wavefront residual is not available directly.

Given then only the Hartmann sensor information and no a-priori information about the wavefront, an acceptable approach is to presume that the null space components in the original wavefront are zero, and then set them to zero in the wavefront reconstruction. There is a practical difficulty however in defining

---

\* Lawrence Livermore National Laboratory, P.O. Box 808, L-290, Livermore, CA 94550

what exactly is the null space. The null property is a matter of degree; for example, on large AO systems such as Keck's where there are many hundreds of sample grid points, parasitic waffle-like behavior can occur over large sections of the aperture, but not necessarily over the entire aperture. These localized waffle shapes are in some sense close to, but not precisely in, the null space, so their magnitudes must become quite large before the Hartmann sensor can see them above the noise. But since these shapes are allowed as a solution, noise in the wavefront measurement may drive them to unacceptably large values.

Once a mode space is defined, one can rank the modes by order of their "measurability" by the Hartmann sensor. Mode measurability is defined as the norm of the Hartmann sensor response vector to the unit norm wavefront mode. In effect, it is the distance from null space. Low-ranked modes could then be arbitrarily rejected from the reconstruction. Now the problem is one of suitably defining such a mode space where relatively measurable modes separate neatly from nearly null modes. This is not always easy. For example the obvious and direct singular value decomposition<sup>4,7</sup> unfortunately results a mode space where *almost all* of the modes have localized waffle-like behavior, so it is not necessarily a good space to use to sort out waffle.

In this paper we describe and analyze the behavior of least-squares reconstructors, particularly with respect to their affect on low-measurability modes. We then introduce a new technique that is successful in reducing anomalous low-measurability behavior such as localized waffle in the wavefront solution. This technique can be adapted so as to remove any specific modes that are undesirable (such as piston, tip, and tilt for example) as well as suppress (the more nebulously defined) partial waffle behavior.

## 2. THE LEAST-SQUARE WAVEFRONT RECONSTRUCTOR

The Hartmann-Shack wavefront sensor can be modeled as the matrix-vector operation

$$\mathbf{s} = \mathbf{H}\mathbf{a} \quad (1)$$

where  $\mathbf{s}$  is an  $n_s$ -vector of Hartmann sensor readings, and  $\mathbf{a}$  is an  $n_a$  vector of wavefront samples (or, without loss of generality,  $\mathbf{a}$  is vector of deformable mirror actuator commands in an adaptive optics system -- from here on we will refer to  $\mathbf{a}$  as the "actuator vector" and the elements of  $\mathbf{a}$  as the "actuator commands").  $\mathbf{H}$  is an  $n_s \times n_a$  actuator influence matrix. The columns of  $\mathbf{H}$  are the Hartmann sensor responses to unit actuator commands.

The  $\mathbf{H}$  matrix for a particular system will depend on the functional form of the actuator influence function and on the aperture boundary. Typically, the number of Hartmann sensor measurements (counting two per subaperture, one for each of the x and y components of slope) exceeds the number of actuators, that is,  $n_s > n_a$ , so the inverse problem implied by (1), find  $\mathbf{a}$  given  $\mathbf{s}$ , is overdetermined. Let  $\hat{\mathbf{a}}$  represent the estimated actuator values. We wish to minimize the sum of squares of the measurement residuals:

$$\tilde{\mathbf{s}}^T \tilde{\mathbf{s}} = (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}})^T (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}}) \quad (2)$$

Taking the partial derivative with respect to  $\hat{\mathbf{a}}$  and setting equal to zero, we come up with the following *normal equations*:

$$\mathbf{H}^T \mathbf{H} \hat{\mathbf{a}} = \mathbf{H}^T \mathbf{s} \quad (3)$$

which can be solved explicitly for  $\hat{\mathbf{a}}$  provided the matrix  $\mathbf{H}^T \mathbf{H}$  is non singular. The matrix  $(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T = \mathbf{H}^*$  is known as the *Moore-Penrose pseudo-inverse*<sup>8</sup> of  $\mathbf{H}$  and has the obvious property that  $\mathbf{H}^* \mathbf{H} = \mathbf{I}$ .

Here we must point out an important fact that is particularly relevant to Hartmann sensors: the matrix  $\mathbf{H}^T \mathbf{H}$  is singular if any null modes exist. The rank deficiency of  $\mathbf{H}^T \mathbf{H}$  is equal to the dimension of the null space of  $\mathbf{H}$ . The Fried geometry, for example, has two null modes: piston and waffle, so the rank deficiency of  $\mathbf{H}$  is 2, and thus solutions to (3) are not explicitly defined via the pseudo-inverse.

## 3. THE SINGULAR VALUE DECOMPOSITION

The existence of a null space of  $\mathbf{H}$  implies there are combinations of actuator command vectors (those in the null space) that will have no effect on the residual square error given by equation (2). It also implies that there are  $n_0$  linear combinations of the columns of  $\mathbf{H}$  that sum to zero. One can imagine building a new

matrix using  $n_a - n_0$  linear combinations of columns of  $\mathbf{H}$  to produce a "thinner"  $n_s \times (n_a - n_0)$  matrix  $\mathbf{H}'$  which is full rank and maps the of actuator vectors not in the null space to the Hartmann measurement space. The pseudo inverse of  $\mathbf{H}'$  ( $=\mathbf{H}'^*$ ) maps sensor vectors to the non-null space of actuators. This method is equivalent to finding, among all the actuator vectors that exactly satisfy the minimization criterion (2), the one of minimum norm, and the one having no projection into the null space.

The mechanics of forming the pseudo-inverse as described above may be accomplished using a numerical method known as the singular value decomposition<sup>7</sup>. The matrix  $\mathbf{H}$  can always be written as the product

$$\mathbf{H} = \mathbf{B} \mathbf{\Sigma} \mathbf{A}^T \quad (4)$$

where  $\mathbf{A}$  is an  $n_a \times n_a$  matrix with orthonormal columns (and rows), that is,  $\mathbf{A}^T = \mathbf{A}^{-1}$ ,  $\mathbf{B}$  is an  $n_s \times n_a$  matrix with orthonormal columns ( $\mathbf{B}^T \mathbf{B} = \mathbf{I}$ ), and  $\mathbf{\Sigma} = \text{diag}(\sigma_i)$  is a diagonal matrix with nonnegative diagonal elements. The columns of  $\mathbf{A}$  define the modes in the actuator space and the columns of  $\mathbf{B}$  describe the normalized sensor response vectors to each of the modes. The quantity  $\sigma_i$  is the strength of the sensor response to actuator mode  $i$ , essentially the mode measurability. Those modes that correspond to  $\sigma_i = 0$  are in the null space. The singular value decomposition is unique in that the response vectors to orthonormal actuator-space vectors are themselves orthogonal in the sensor space. The singular value decomposition (4) can be obtained numerically using (for example) the SVDC routine in IDL.

We can write the matrices  $\mathbf{A} = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_{na}]$  and  $\mathbf{B} = [\mathbf{b}_1 | \mathbf{b}_2 | \dots | \mathbf{b}_{na}]$ . Substituting (4) into (1) we can write the sensor vector as a sum of modal contributions:

$$\mathbf{s} = \sum_i \mathbf{b}_i \sigma_i (\mathbf{a}_i \bullet \mathbf{a}) \quad (5)$$

This can be interpreted as follows: find the  $i$ 'th mode coefficient of the actuator vector  $\mathbf{a}$  by taking the inner product with mode vector  $\mathbf{a}_i$ , multiply it by  $\sigma_i$ , then make this the  $i$ 'th mode coefficient for the sensor vector using mode  $\mathbf{b}_i$  in sensor space. The index  $i$  in the sum might as well cover only those modes not in the null space, since null space modes have  $\sigma_i = 0$ . So if we take  $\mathbf{A}' =$  columns of  $\mathbf{A}$  not in the null space,  $\mathbf{B}' =$  corresponding columns of  $\mathbf{B}$  consisting of  $\mathbf{b}_i$  where  $\mathbf{a}_i$  is not in the null space, and  $\mathbf{\Sigma}' = \text{diag}(\sigma_i \neq 0)$  then

$$\mathbf{H} = \mathbf{B}' \mathbf{\Sigma}' \mathbf{A}'^T \quad (6)$$

and  $\mathbf{B}'^T \mathbf{B}' = \mathbf{I}$  and  $\mathbf{A}'^T \mathbf{A}' = \mathbf{I}$ , so  $\mathbf{H}^T \mathbf{H} = \mathbf{A}' \mathbf{\Sigma}'^2 \mathbf{A}'^T$ . It is easy to verify by direct substitution that

$$\hat{\mathbf{a}} = \mathbf{A}' \mathbf{\Sigma}'^{-1} \mathbf{B}'^T \mathbf{s} \quad (7)$$

satisfies the normal equations (3). Furthermore, the solution is the one of minimum norm since  $\hat{\mathbf{a}} \bullet \mathbf{a}_i = 0$  for all  $\mathbf{a}_i$  in the null space, which is verifiable using (7) and the fact that the columns of  $\mathbf{A}'$  consist only of vectors not in the null space. The solution can be written as a sum of modal contributions:

$$\hat{\mathbf{a}} = \sum_{i: \sigma_i \neq 0} \mathbf{a}_i \sigma_i^{-1} (\mathbf{b}_i \bullet \mathbf{s}) \quad (8)$$

which can be interpreted as follows: determine the  $i$ 'th mode coefficient for the sensor vector by taking the inner product with  $\mathbf{b}_i$ , multiply by the *inverse* of the mode's measurability,  $\sigma_i$ , and then use this as the coefficient for the  $i$ 'th actuator mode in the solution. The sum is taken only over the modes not in the null space ( $\sigma_i \neq 0$ ).

Equation (8) shows the inherent difficulty in using the singular value decomposition to implement the wavefront reconstructor. The modes with the least (but nonzero) measurability get amplified the most in the reconstruction, making them prone to noise in  $\mathbf{s}$  or error in the determination of  $\mathbf{b}_i$  (induced by errors in determining  $\mathbf{H}$ ). Furthermore, "throwing out" modes with low measurability (say by setting a threshold on  $\sigma_i$ ) does not guarantee that the remaining modes do not have anomalous behavior, such as waffle over portions of the aperture (Figures 1-3).

## 4. WEIGHTED LEAST SQUARES

A variation along the least-squares theme is the weighted-least-squares criterion

$$\xi^T \mathbf{W} \xi = (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}})^T \mathbf{W} (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}}) \quad (9)$$

where  $\mathbf{W}$  is a positive-definite matrix.  $\mathbf{W} = \text{diag}(w_i)$  for example could be used to indicate relative confidence in the individual sensor readings. The normal equation is

$$\mathbf{H}^T \mathbf{W} \mathbf{H} \hat{\mathbf{a}} = \mathbf{H}^T \mathbf{W} \mathbf{s} \quad (10)$$

Substituting (6) into (10), and solving for  $\hat{\mathbf{a}}$  we get

$$\hat{\mathbf{a}} = \mathbf{A}' \Sigma'^{-1} (\mathbf{B}'^T \mathbf{W} \mathbf{B}')^{-1} \mathbf{B}'^T \mathbf{W} \mathbf{s} \quad (11)$$

noting that the indicated matrix inverse exists since  $\mathbf{B}'$  is full rank and  $\mathbf{W}$  is positive definite. We can make the following observations. The mode space is defined by the same  $\mathbf{a}_i$  (columns of  $\mathbf{A}$ ) as in the non weighted least squares case, so the mode space is unaffected by the weighting.  $\mathbf{W}^{1/2} \mathbf{H}$  has the same null space as  $\mathbf{H}$ , therefore weighting the residuals introduces no penalty for anomalous behavior, that is, arbitrary vectors in null space can be added to  $\hat{\mathbf{a}}$  and still satisfy (10). (11) is the minimum norm solution (no components in the null space) however, aside from the fact that the measurements are post multiplied by a fancier conditioning matrix  $(\mathbf{B}'^T \mathbf{W} \mathbf{B}')^{-1} \mathbf{B}'^T \mathbf{W}$  to determine mode coefficients, the low measurability modes are still experience the highest amplification via  $\Sigma'^{-1}$ . We could imagine that some adjustment of  $\mathbf{W}$  might cause this conditioning matrix to suppress low measurability modes, but this is not the case as is shown below.

If the measurement vector  $\mathbf{s}$  consists only of components that are mapped by  $\mathbf{H}$  from the non-null space (that is,  $\mathbf{s}$  is noise free and  $\mathbf{H}$  is exact, as in equation (1)), then  $\mathbf{s}$  is a linear combination of the columns of  $\mathbf{H}$ . From (6) it is clear that  $\mathbf{s}$  is a linear combination of the columns of  $\mathbf{B}'$  as well, that is

$$\mathbf{s} = \mathbf{B}' \mathbf{c} \quad (12)$$

where  $\mathbf{c}$  is an  $n_r - n_0$  vector of modal coefficients:  $c_i = \mathbf{b}_i \bullet \mathbf{s}$ . Substituting (12) into (11) yields

$$\hat{\mathbf{a}} = \mathbf{A}' \Sigma'^{-1} \mathbf{c} \quad (13)$$

which is *independent of the weighting*. So weighting of the measurement vector does not at all affect the solution's modal behavior, and in particular, low measurability modes (such as partial waffle) will still be highly amplified in exactly the same way as (non-weighted) least squares. We should note in all fairness to the efficacy of weighted-least-squares however, that the *noise amplification* will be quite different depending on  $\mathbf{W}$ . With measurement noise, the statistical least mean-square error estimate of  $\mathbf{a}$  will be obtained by setting  $\mathbf{W} = \mathbf{V}^{-1}$  where  $\mathbf{V}$  is the covariance of the measurement noise.

## 5. REGULARIZATION

Another approach is to force the explicit solution of (3) by replacing  $\mathbf{H}^T \mathbf{H}$  by  $\mathbf{H}^T \mathbf{H} + \alpha^2 \mathbf{I}$  where  $\alpha$  is a positive constant called the *regularization parameter*. The matrix  $\mathbf{H}^T \mathbf{H} + \alpha^2 \mathbf{I}$  is full rank and therefore invertible, so we create the regularized pseudo-inverse

$$\mathbf{H}_r^* = (\mathbf{H}^T \mathbf{H} + \alpha^2 \mathbf{I})^{-1} \mathbf{H}^T \quad (11)$$

An advantage of this approach is that it does not require a singular value decomposition to find and remove the null space. A disadvantage is that  $\mathbf{H}_r^* \mathbf{H} \neq \mathbf{I}$ . It is interesting to note that  $\hat{\mathbf{a}} = \mathbf{H}_r^* \mathbf{s}$  is the explicit minimization of the modified least-squares criterion

$$J = (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}})^T (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}}) + \alpha^2 \hat{\mathbf{a}}^T \hat{\mathbf{a}} \quad (12)$$

which can be verified by taking the derivative of  $J$  with respect to  $\hat{\mathbf{a}}$  and setting equal to zero. The modified criterion places a penalty on the use of actuator "energy" that offsets the penalty due to measurement fitting error.

The normal equation is

$$\left(\mathbf{H}^T \mathbf{H} + \alpha^2 \mathbf{I}\right) \hat{\mathbf{a}} = \mathbf{H}^T \mathbf{s}. \quad (13)$$

Since

$$\left(\mathbf{A} \Sigma^2 \mathbf{A}^T + \alpha^2 \mathbf{I}\right) = \mathbf{A} \left(\Sigma^2 + \alpha^2 \mathbf{I}\right) \mathbf{A}^T, \quad (14)$$

we see that the regularization is simply equivalent to replacing the singular values by weighted versions of them, creating a "soft lower limit" on each singular value:

$$\sigma_i^{-2} \rightarrow \left( \frac{1}{\sigma_i^2 + \alpha^2} \right) \quad (15)$$

but that the singular mode space vectors are unchanged. The only effect of regularization therefore is to suppress amplification of low-singular value modes in the reconstructor by asymptotically limiting  $\sigma_i^{-2}$  to a maximum value of  $\alpha^2$ . This is a reasonable alternative to throwing out the low singular value modes completely as was suggested for the non-regularized case, although it is basically the same idea. It still does not solve the problem of localized waffle being present in every mode since the mode space remains unchanged.

## 6. ACTUATOR PENALTY

We now present a weighting approach that will actually change the mode space structure. We introduce an actuator weighting, but instead of using a scalar regularization parameter we use a positive-definite matrix,  $\mathbf{V}$ .

$$J = (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}})^T (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}}) + \hat{\mathbf{a}}^T \mathbf{V} \hat{\mathbf{a}} \quad (16)$$

For example, if  $\mathbf{V}$  were the outer product of actuator waffle mode, then the added term will penalize the waffle mode. But, actually, we want to penalize all local waffle behavior as well. To do this, we introduce the weighting matrix

$$\mathbf{V} = \mathbf{F}^T \mathbf{F} \quad (17)$$

where  $\mathbf{F}$  is described as:

$$\mathbf{F} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \otimes \quad (18)$$

that is,  $\mathbf{F}$  is a matrix form that implements convolution with a 2x2 localized waffle. Note that instead of  $\mathbf{V}$  being a rank-1 matrix that penalizes only one mode,  $\mathbf{V}$  is a rank N matrix that penalizes all localized waffle behavior. We can add rank-1 matrices to  $\mathbf{V}$  as well, for example to penalize pure piston mode. The reconstruction matrix is now

$$\mathbf{K} = \left(\mathbf{H}^T \mathbf{H} + \mathbf{F}^T \mathbf{F}\right)^{-1} \mathbf{H}^T \quad (19)$$

The singular value decomposition of  $\mathbf{K}$  can be written:

$$\mathbf{K} = \mathbf{A}_K \Sigma_K \mathbf{B}_K^T \quad (20)$$

Since  $\mathbf{K}$  is the reconstructor, it is the larger values of  $\sigma_K$  that correspond to the modes that are most amplified. Examining the mode space of this new reconstructor (Figures 4-6), we see that this technique acts to sort the modes in the way we want. Those modes with large singular values (greatest amplification)

have the appearance of low-order Zernike behavior, with no waffle. The modes with the small singular values have localized waffle, and there is a general correspondence between the magnitude of the singular value and apparent amount of localized waffle in the mode.

## 7. STATISTICAL WEIGHTING

We might mention at this point that a-priori statistical information concerning the quantities to be determined (the phases) and the measurement noise can be used to set the weighting matrices in the cost function terms. The general form is

$$J = (\mathbf{s} - \mathbf{H}\hat{\mathbf{a}})^T \mathbf{W}(\mathbf{s} - \mathbf{H}\hat{\mathbf{a}}) + \hat{\mathbf{a}}^T \mathbf{V}\hat{\mathbf{a}} \quad (21)$$

If we set  $\mathbf{V}$  equal to the inverse of the a-priori covariance matrix for phase points and  $\mathbf{W}$  equal to the inverse of the measurement noise covariance matrix, and the statistics are Gaussian, then (we state here without proof) minimizing  $J$  is equivalent to finding the statistically most likely (conditional mean) value of actuator vector  $\mathbf{a}$  given the measurement vector  $\mathbf{s}$ . For example, the initial uncertainty of atmospheric phase is given by Kolmogorov statistics. The  $k^{-11/3}$  spatial power spectrum encoded in  $\mathbf{V}$  will act to increase the penalty for high spatial frequency behavior in the solution, since the atmosphere is unlikely to contain high spatial frequency components relative to low spatial frequency ones. This acts fortuitously to suppress waffle behavior, which is a high spatial frequency behavior.

## 8. CONCLUSION

We have presented analyses of various weighting schemes for least squares wavefront reconstructors where the actuators and Hartmann subapertures are arranged on a rectilinear grid. A number of the more commonly used methods are susceptible to anomalous behavior, particularly the amplification of localized waffle. Furthermore, the singular value decomposition of the reconstructor matrix typically produces a mode set where almost all of the modes appear to contain some amount of localized waffle, frustrating attempts to filter out waffle modes by selection. We found that only a modification of the cost function by including a penalty term on the actuators will actually modify the mode space. With a suitable weighting matrix on the actuators, waffle behavior can be sorted out into modes that have the least influence in the reconstructor matrix, thus suppressing waffle in the wavefront solutions.

## ACKNOWLEDGEMENT

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48. This work has been supported by the National Science Foundation Science and Technology Center for Adaptive Optics, managed by the University of California at Santa Cruz under cooperative agreement No. AST - 9876783.

## REFERENCES

1. Herrmann, J., *Least-squares wave front errors of minimum norm*, **J. Opt. Soc. Am.**, Vol 70, No. 1, (1980), 28-35.
2. Southwell, W. H., *Wave-front estimation from wave-front slope measurements*, **J. Opt. Soc. Am.**, Vol. 70, No. 8, (1980), 998-1006.
3. Fried, D. L., *Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements*, **J. Opt. Soc. Am.**, Vol 67, (1977), 370-375.
4. Hardy, J. W., **Adaptive Optics for Astronomical Telescopes**, Chapter 8, Oxford University Press, (1998).
5. Brase, J. M., An, J., Avicola, K., Beeman, B. V., Gavel, D. T., Hurd, R., Johnston, b., Jones, H., Kuklo, T., Mac, C. E., Olivier, S. S., Waltjen, K. E., Watson, J., *The wavefront control system for the Keck Telescope*, **Proc. SPIE**, Vol. 3352, (1998), 517-521.
6. Born, M., and Wolf, E., **Principles of Optics**, Pergamon Press, Oxford, (1975).
7. Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T., **Numerical Recipes**, Cambridge University Press, Cambridge, (1986).
8. Penrose, R. A., *Generalized Inverse for Matrices*, **Proc. Cambridge Phil. Soc.**, 51, (1955), 406-413.

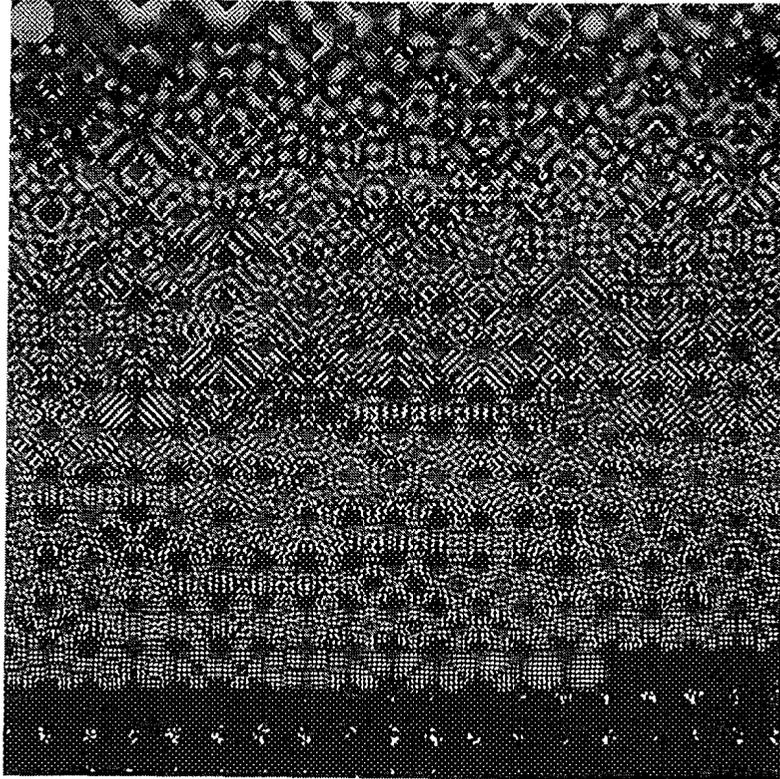


Figure 1. The mode set for the singular value decomposition (4) for an  $18 \times 18$  grid of actuators on a circular aperture, ranked in order of magnitude of the singular values of  $H$  (smallest is upper left). All the modes show anomalous waffle behavior.

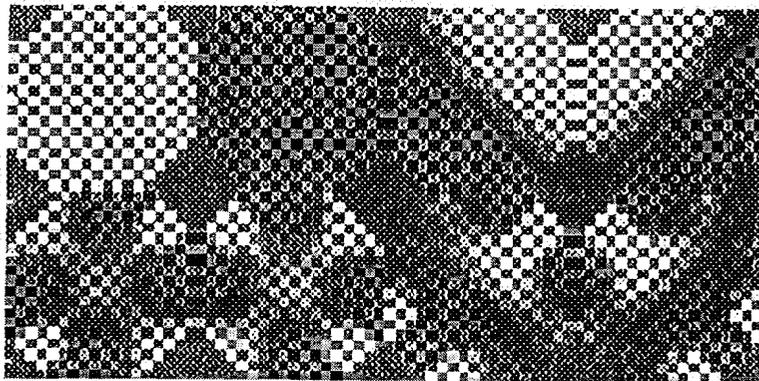


Figure 2. Zoomed in view of modes 1-4 (top row) and 19-22 (bottom row). Note that modes 3 and 4 are “waffle/anti-waffle” over opposite halves of the aperture and are therefore orthogonal to the “pure” waffle. Such modes slip past mode-rejection algorithms based on dot-products with the pure waffle.

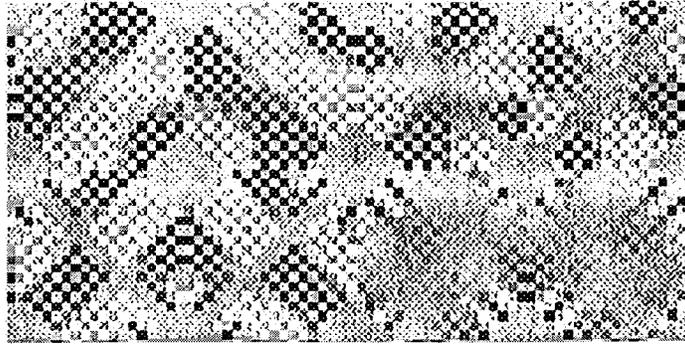


Figure 3. Zoomed in view of other modes 15-18 (top row) and 33-36 (bottom row) in the SVD mode set. Every mode shows waffle behavior.

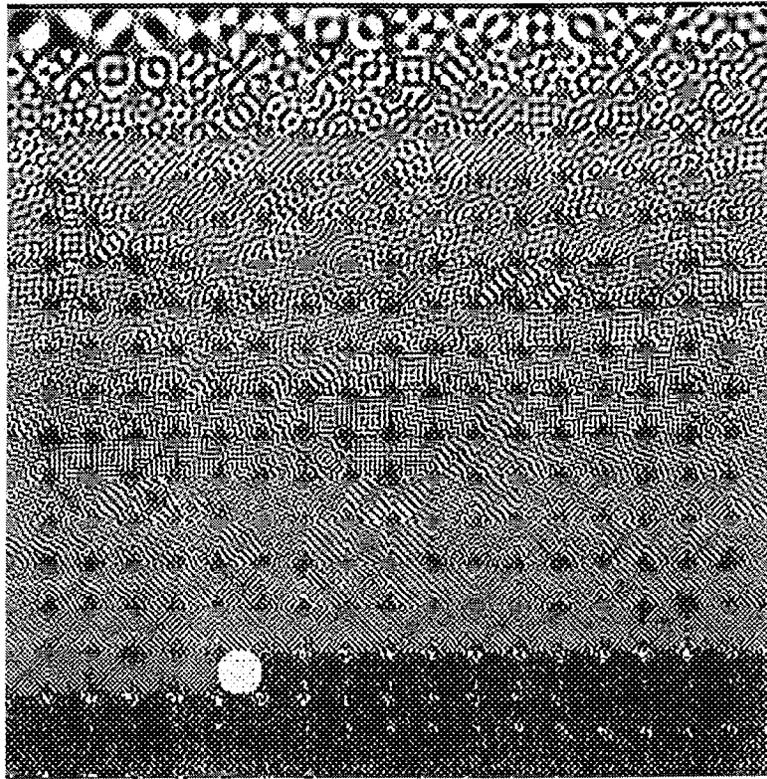


Figure 4. The mode set resulting from the actuator-penalty method (20), ranked in reverse order of magnitude of the singular value of the reconstruction matrix  $\mathbf{K}$  (largest is upper left). Waffle behavior is relegated to modes with the lowest singular values. Modes with high singular values look like low-order Zernikes, with no waffle. Piston is the lowest non-zero singular value mode (by design – we added piston-mode actuator penalty). The “zero-visibility” (zero singular value) modes are associated with actuators outside the aperture and so are last in this set.



Figure 5. Zoomed in view of modes 1-4 (top row) and 19-22 (bottom row) from the actuator-penalty method.

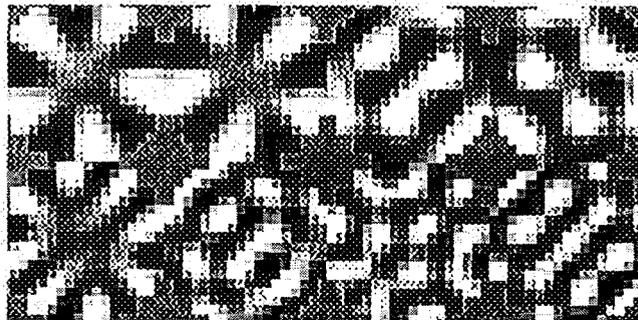


Figure 6. Zoomed in view of modes 15-18 (top row) and 33-36 (bottom row) from the actuator-penalty method. Although high-order modes have high spatial frequencies, they have no localized waffle behavior.



Figure 7. Long-exposure point-spread-functions from a closed loop simulation using a) a standard SVD pseudo-inverse matrix, with modes rejected if the singular value falls below 0.15 times the maximum singular value, and b) a reconstruction matrix generated using the actuator-penalty method with no modes rejected. The deformable mirror consists of a circular aperture with 276 actuators arranged on a rectilinear grid (within the circle inscribing a square 18x18 grid). The input wavefront presented to the controller was flat (zero phase over the aperture for the duration of the simulation) while the Hartmann slope measurement error was 1 radian per subaperture-width rms white noise, spatially and temporally uncorrelated. The characteristic waffle-dot far-field pattern is quite evident in image a, while there is no hint of it in image b. Waffle bumps are about 1/10 the amplitude of the peak in image a. (Images are shown in log-stretched greyscale.) The Strehl in image b is 0.66 compared to 0.42 in image a.