

**Efficient Computation of Minimum Variance Wavefront
Reconstructors Using Sparse Matrix Techniques**

Brent L. Ellerbroek^a

Gemini Preprint #82

a. Gemini Observatory, Hilo, HI 96720, USA

Efficient computation of minimum variance wavefront reconstructors using sparse matrix techniques

Brent L. Ellerbroek

Gemini Observatory, Hilo, HI, USA

bellerbroek@gemini.edu

The complexity of computing conventional matrix multiply wavefront reconstructors scales as $O(n^3)$ for most adaptive optical (AO) systems, where n is the number of deformable mirror (DM) actuators. This is impractical for proposed systems with extremely large n . It is known that sparse matrix methods improve this scaling for least squares reconstructors, but sparse techniques are not immediately applicable to the minimum variance reconstructors now favored for multi-conjugate adaptive optics (MCAO) systems with multiple wave front sensors (WFS's) and DM's. Complications arise from the non-sparse statistics of atmospheric turbulence, and the global tip-tilt WFS measurement errors associated with laser guide star (LGS) position uncertainty. In this paper, we describe how sparse matrix methods can still be applied by use of a sparse approximation for turbulence statistics, and by recognizing that the non-sparse matrix terms arising from LGS position uncertainty are low rank adjustments that can be evaluated using the matrix inversion lemma. Sample numerical results for AO and MCAO systems illustrate that: The approximation made to turbulence statistics has negligible effect on estimation accuracy, the time to compute the sparse minimum variance reconstructor for a conventional natural guide star (NGS) AO system scales as $O(n^{3/2})$ and is only a few seconds for $n = 3500$, and sparse techniques reduce the reconstructor computations by a factor of 8 for sample MCAO systems with 2417 DM actuators and 4280 WFS subapertures. Extrapolating to 9700 actuators and 17120 subapertures, we predict a reduction by a factor of about 30 or 40 to 1. © 2002 Optical Society of America

OCIS codes: 010.1080

1 Introduction

Adaptive optical (AO) systems are used in ground-based optical and near-IR astronomical telescopes to correct for the phase aberrations induced by atmospheric turbulence [1, 2, 3]. These aberrations are compensated by adjusting the figure of a deformable mirror (DM) to null the residual phase errors as measured by a wave front sensor (WFS). The control algorithm used to determine the DM commands from the WFS measurements is frequently referred to as the wavefront reconstructor. Analytical and numerical methods for optimizing and evaluating reconstructors are very well developed, both for existing conventional AO systems [4, 5, 6, 7], proposed AO systems with multiple laser guide stars (LGS's) [8, 9], and proposed multi-conjugate AO (MCAO) systems that would employ multiple DM's and WFS's to compensate for atmospheric turbulence across an extended field-of-view [10, 11, 12, 13, 14]. In comparison, work on efficiently computing and implementing reconstruction algorithms is at a less advanced state. Since wavefront reconstruction is a linear estimation problem for the case of astronomical AO, the DM actuator command vector a can be determined from the WFS measurement vector s using a matrix multiply. The complexity of computing the reconstruction matrix using standard methods scales as $O(n^3)$, where n is the dimensionality of a . This will become impractical for proposed MCAO and "Extreme" AO (ExAO) systems on extremely large telescopes (ELT's), where n may be on the order of 10,000 or even 100,000.

Fortunately, it has been known for some time that sparse matrix techniques may be used to significantly reduce the numerical complexity of classical least squares reconstruction algorithms, which determine a by minimizing the RMS value (or L^2 norm) of the residual measurement of s [15]. This method exploits the fact that the DM-to-WFS influence matrix is sparse for many AO component technologies, so that an adjustment to a single DM actuator effectively influences WFS measurements from only a few neighboring subapertures. The essential reason this reduces computational complexity is the fact that a sparse, banded matrix M may be factored as $M = LL^T$ with L sparse and lower triangular, so that the system $x = M^{-1}y$ can be solved for x in terms of y by first solving $Lv = y$ and then $L^T x = v$ by back substitution. Sparse matrix reconstructors and other efficient approaches such as Fourier transform reconstructors, multi-grid methods [16], and preconditioned conjugate gradient algorithms [17] are promising approaches for future ExAO systems, since theory and simulation indicate that conventional least squares reconstruction is near-optimal for this class of AO [18]. Sparse matrix methods have also been proposed for efficiently deconvolving a known, but spatially varying, point spread function from a blurry image [19].

However, sparse matrix methods are not immediately applicable to reconstruction algorithms for MCAO systems for several reasons. Conventional least squares reconstructors generally perform poorly for MCAO, and a regularized, or modally filtered, algorithm is necessary to obtain optimal or near-optimal performance [20]. Two more satisfactory approaches are SVD filtering of the reconstructor to suppress poorly sensed modes and minimum variance wavefront reconstruction [7], but both techniques yield full matrices that are incompatible with the direct application of sparse matrix techniques. Secondly, the position uncertainty problem for laser guide stars introduces high levels of full aperture tip/tilt-measurement noise in LGS WFS measurements. The global tip/tilt terms must be filtered from the LGS measurements, and auxiliary tip/tilt natural guide star (NGS) WFS's must be included in the MCAO system to measure the full-aperture wavefront tilt. Both of these effects complicate the structure of the DM-to-WFS influence matrix, with the result that sparse matrix methods once again cannot be immediately applied.

This paper describes techniques for dealing with these difficulties to obtain a representation for the minimum variance wavefront reconstructor that can be efficiently evaluated using sparse matrix methods, even for the case of a MCAO system. First, the non-sparse regularization term appearing

in the reconstructor is replaced by a sparse approximation. In a heuristic sense this approximation is equivalent to replacing the Kolmogorov $\kappa^{-11/3}$ spatial power spectrum for atmospheric turbulence (where κ is a spatial frequency variable) with κ^{-4} . It may be a surprise that this approximation has only a negligible impact upon the performance of the reconstructor, increasing the mean-square residual phase error by from 0.1% to no more than 1.5% for all of the cases we have evaluated. Secondly, the non-sparse matrix terms appearing in the minimum variance reconstructor due to LGS position uncertainty and the inclusion of NGS tip/tilt WFS's in the AO system are of low rank, which allows sparse matrix methods to still be applied with the aid of the matrix inversion lemma. Briefly, this lemma implies that if U and V are matrices with only a few columns and are dimensioned such that $M + UV^T$ is defined, then $(M + UV^T)^{-1}$ is equal to $M^{-1} + U'V'^T$, where the matrices U' and V' have the same dimensions as U and V . It follows that if M has a structure that allows the system $x = M^{-1}y$ to be solved efficiently (*e.g.*, a sparse matrix), the system $x = (M + UV^T)^{-1}y$ can be solved efficiently as well.

For sample cases involving NGS and LGS MCAO systems with about 2400 total DM actuators and 4280 WFS subapertures on a 16 meter telescope aperture, the above methods provide about a factor of 8 improvement in the time needed to compute the minimum variance reconstructor. Extrapolating to a 32 meter MCAO system with about 9700 actuators and 17120 subapertures we predict an improvement by a factor of about 30 or 40, although at present our computer does not have sufficient memory to evaluate this case. For ExAO the computation requirements for this sparse implementation of the minimum variance reconstructor are comparable to results obtained previously for conventional least squares reconstruction, allowing reconstructors to be computed in only a few seconds for very high-order AO systems.

The remainder of this paper is organized as follows. Section 2 reviews prior results on the use of sparse matrix techniques for least squares reconstructors. Section 3 derives ones of the standard formulas for the minimum variance reconstructor, which decomposes into two parts: Estimating the full turbulence profile, and then finding the DM actuator commands that best correct for this profile over the desired field-of-view [13]. The sparse matrix methods for efficiently evaluating these two operators are presented in Sections 4 and 5. Section 6 summarizes sample simulation results obtained on reconstructor performance and computation requirements for a range of conventional AO, ExAO, and MCAO system configurations.

2 Least Squares Wavefront Reconstruction

In this section, we review the application of sparse matrix techniques to classical least squares wavefront reconstruction, as first presented in [15]. The improvements that can be obtained in computational efficiency are very impressive, at least for the case of a conventional AO system with a single DM and a single NGS WFS.

Classical least squares wavefront reconstruction is based upon the WFS measurement model

$$s = G_a a + n. \tag{1}$$

Here s is the WFS measurement vector, a is the DM actuator vector to be estimated from s , $G_a = \partial s / \partial a$ is the DM-to-WFS influence matrix, and n is additive, zero mean measurement noise. All components of n are assumed to be uncorrelated and of equal variance. Note that atmospheric turbulence is not included in this measurement model, and the goal of least squares wavefront reconstruction is simply to determine an estimate \hat{a} of a that yields the best mean-square fit to s in Eq. (1). The value of \hat{a} is given formally by the expression

$$\hat{a} = \arg \min_a \|s - G_a a\|^2, \tag{2}$$

and this minimization problem can be solved by determining the value of a for which the partial derivatives $\partial||s - G_a a||^2/\partial a$ are identically zero. This is a standard linear least squares problem, since the merit function to be minimized is quadratic in a . The minimum norm solution is given by the formula [4, 5, 6]

$$\hat{a} = (G_a^T G_a)^\dagger G_a^T s, \quad (3)$$

where the superscript “ T ” denotes the transpose of a matrix or vector, and M^\dagger is the pseudo-inverse of the matrix M . Although this is nearly the most elementary reconstruction algorithm possible, least squares estimation has been applied very successfully in many hardware systems and remains perhaps the most commonly used AO control algorithm even today.

The classical least squares reconstruction algorithm is of complexity $O(n_a^3)$ to compute and $O(n_a^2)$ to apply when implemented using conventional matrix inversions and matrix/vector multiplies, where n_a is the dimension of the DM actuator command vector a [21]. This becomes a significant practical limitation for values of n_a in excess of about 1000, let alone the values of 10,000 or even 100,000 presently under consideration for so-called “extreme AO” systems. Fortunately the matrix G_a is highly sparse for many WFS and DM technologies, since each individual DM actuator influence function couples into only a few (perhaps 4 or 8) elements of the wavefront sensor measurement vector s . It follows that the matrix $G_a^T G_a$ is highly sparse as well, and this matrix can be decomposed into a Cholesky factorization

$$G_a^T G_a = LL^T, \quad (4)$$

where L is lower triangular and also sparse. In fact, the sparseness of L can be improved by reordering the columns of the influence matrix G_a , which corresponds to simply renumbering the elements of the DM actuator command vector [22, 23]. Once the matrix $G_a^T G_a$ has been factored, \hat{a} may be computed from s in three steps using the equations

$$v = G_a^T s; \quad Lw = v; \quad L^T \hat{a} = w. \quad (5)$$

The intermediate variable v is first computed from s using the first equation, the second equation is next solved for w using backsubstitution, and the third equation is then similarly solved for \hat{a} . Each of the matrices G_a^T , L , and L^T is sparse, leading to a very significant reduction in computational complexity. The reader should note that this technique still obtains the exact least squares solution, and is in fact less subject to roundoff error than the conventional matrix multiply solution when n_a is very large.

The most dramatic reduction in computation requirements achieved through the use of sparse matrix techniques are obtained not in applying the least squares wavefront reconstructor, however, but in initially computing the matrix factorization $G_a^T G_a = LL^T$. Fig. 1 plots the number of floating point operations required for this decomposition as a function of n_a for a square aperture geometry and the so-called “Hudgins” or “shearing interferometer” wavefront sensor subaperture geometry [5]. The number of computations needed to explicitly compute $(G_a^T G_a)^\dagger$ is also plotted for comparison. The computational requirements for the two approaches scale as $n_a^{3/2}$ and n_a^3 , respectively, leading to a factor of 3.35×10^5 advantage for the sparse matrix methods for the case of an extreme AO system with $n_a = 10^4$. This is a very nontrivial reduction in computational complexity, but unfortunately the classical least squares reconstructor is not always an appropriate control algorithm for more sophisticated AO system configurations. A more general reconstruction algorithm that can be optimized for these applications is described below.

3 Minimum Variance Wavefront Reconstruction

The standard least squares wavefront reconstruction algorithm performs poorly for some AO applications because the WFS measurement model described by Eq. (1) is oversimplified. The WFS measurement vector s is in fact a function of the atmospheric turbulence profile, not a set of pre-existing DM actuator command errors that must be nulled to obtain a perfect wavefront. The DM actuator command vector a should be selected to compensate the turbulence-induced wavefront error, which is not in general the same problem as finding the best fit to the WFS measurement vector s . If the statistics of the atmospheric turbulence profile and the WFS measurement noise are known, minimum variance wavefront reconstruction [7] provides an optimal solution in the sense of (as the name implies) minimizing the variance of the residual wavefront error remaining after the DM actuator commands have been applied. This section briefly derives one of the standard representations for the minimum variance reconstructor that can be, with some further work, efficiently evaluated using sparse matrix techniques. This representation is general enough that it may be applied to AO applications involving one or several wavefront sensors, deformable mirrors, atmospheric turbulence layers, and wavefronts to be corrected [8, 11].

A Problem Formulation

The residual wavefront profile(s) remaining after commands have been applied to the deformable mirror(s) will be denoted ϕ , and is defined by the equation

$$\phi = H_x x - H_a a. \quad (6)$$

The components of ϕ are the values of the residual phase profile(s) at a set of grid points in the telescope aperture plane, x is a vector of phase values on a grid of points on the atmospheric phase screen(s), a is the DM actuator command vector, and the columns of the matrices H_x and H_a are the “influence functions” associated with the discrete phase points and DM actuators. The effects of x and a on the phase ϕ are assumed to be linear, and are evaluated by tracing rays through the phase screens and DM conjugate planes as illustrated in Fig. 2a. The vector x is a random variable with zero mean and finite second-order statistics. These statistics are typically modeled using the Kolmogorov or von Karman spectrum.

The mean-square residual piston-removed wavefront error will be denoted σ^2 , and is related to ϕ by the formula

$$\sigma^2 = \phi^T W \phi, \quad (7)$$

where W is a symmetric, semi-positive-definite matrix. The coefficients of W may be defined, for example, so that the value of σ^2 is equal to the mean-square piston-removed value of a continuous phase profile obtained by interpolating a smooth function through the values of ϕ specified on the discrete set of grid points [12]. The additional features of W needed to apply sparse matrix methods to this problem are outlined in Section 4 below.

The DM actuator command vector a is computed from the WFS measurement vector s using a linear reconstruction algorithm of the form

$$a = E s, \quad (8)$$

where E is the wavefront reconstruction matrix. The WFS measurement s is modeled as

$$s = G_x x + n, \quad (9)$$

where G_x is the phase-to-WFS influence matrix and n is WFS measurement noise. The elements of G_x are once again computed by tracing rays from the guide star(s) through the phase screen(s)

to the wavefront sensing subapertures, as illustrated in Fig. 2b. Eq. (9) differs from Eq. (1) for the least squares reconstruction algorithm in three important ways: s is now a function of x instead of a , the case of multiple wavefront sensors is allowed, and n is a random vector with zero mean and finite second order statistics. This more general noise model is required, for example, to model multiple natural guide stars of different brightnesses and the effects of LGS position uncertainty.

In this notation, the minimum variance reconstructor E_* is the value of E that minimizes the expected value of σ^2 averaged over the statistics of the phase profile x and the WFS measurement noise n . We generalize this definition slightly and write

$$E_* = \arg \min_E \langle \sigma^2 + k||a||^2 \rangle, \quad (10)$$

where the angle brackets, $\langle \dots \rangle$, denote ensemble averaging over the statistics of noise and turbulence. As shown below, the regularization term $k||a||^2 = ka^T a$ must be included (with a very small value of k) to avoid singularities if the subspace of DM actuator commands having no effect on σ^2 is not *a priori* known.

B Deriving the Reconstructor

Determining the minimum variance reconstructor E_* that minimizes the mean square phase variance σ^2 is a least squares minimization problem in the coefficients of E . Using the definitions for σ^2 and a given above in Eq.'s (7) and (8), we may write

$$\langle \sigma^2 + k||a||^2 \rangle = \langle (H_x x - H_a E s)^T W (H_x x - H_a E s) + k s^T E^T E s \rangle, \quad (11)$$

for the merit function to be minimized. The partial derivatives of this quantity with respect to the coefficients of E must vanish for the optimal value of the reconstructor, *i.e.*

$$0 = \left. \frac{\partial \langle \sigma^2 + k||a||^2 \rangle}{\partial E_{ij}} \right|_{E=E_*}. \quad (12)$$

Differentiating Eq. (11) with respect to the coefficients of E and rearranging yields the result

$$\langle s (H_a^T W H_x x)^T \rangle = \langle s (H_a^T W H_a E_* s)^T \rangle + k \langle s (E_* s)^T \rangle. \quad (13)$$

It is convenient to introduce the notation

$$C_{vw} = \langle v w^T \rangle \quad (14)$$

for the covariance of two zero-mean random variables v and w . By factoring the non-random matrices in Eq. (13) outside of the expected value operations, this expression can be rewritten in the form

$$(H_a^T W H_x) C_{xs} = (H_a^T W H_a + kI) E_* C_{ss}. \quad (15)$$

Eq. (15) can be solved immediately for the minimum variance reconstructor E_* if the matrices C_{ss} and $H_a^T W H_a + kI$ are invertible. The first of these two matrices will be invertible whenever the WFS measurement noise n is nonzero, and the remainder of the paper will restrict attention to this real-world case. The second matrix $H_a^T W H_a$ will generally not be invertible for the case $k = 0$, *i.e.* the formally exact value of the minimum variance reconstructor. This matrix will have a nontrivial null space since there are some modes of DM actuator commands that have no effect on the piston-removed residual wavefront error, for example overall piston adjustments to one or more of the deformable mirrors in the AO system. Clearly, there is no performance penalty if we require

that the output of the wavefront reconstruction matrix be orthogonal to $\text{null}(H_a^T W H_a)$, the null space of $H_a^T W H_a$. This requirement is expressed by the constraint equation

$$N_w^T E_* = 0, \quad (16)$$

where the columns of the matrix N_w are a linearly independent set of vectors from $\text{null}(H_a^T W H_a)$. This constraint implies the condition $N_w N_w^T E_* C_{ss} = 0$ as well, which may be summed with Eq. (15) to yield the result

$$(H_a^T W H_x) C_{xs} = (H_a^T W H_a + N_w N_w^T + kI) E_* C_{ss}, \quad (17)$$

If we are certain that we know a full basis for $\text{null}(H_a^T W H_a)$ it is now possible to set $k = 0$ and compute the precise value of the minimum variance reconstructor. However, this assumption may not hold for MCAO systems due to the richer cross-coupling between the actuators on the multiple deformable mirrors, and in this case k must be set to a (very small) nonzero value to avoid singularities.

[In passing, we note for future use that N_w will be a low rank matrix with only a small number of columns, certainly much smaller than the dimension of the DM actuator command vector a .]

Solving Eq. (17) for the minimum variance reconstructor E_* now yields the result

$$\begin{aligned} E_* &= (H_a^T W H_a + N_w N_w^T + kI)^{-1} (H_a^T W H_x) C_{xs} C_{ss}^{-1} \\ &= F_x E_x, \end{aligned} \quad (18)$$

where the variables F_x and E_x are abbreviations for the terms

$$F_x = (H_a^T W H_a + N_w N_w^T + kI)^{-1} (H_a^T W H_x) \quad (19)$$

$$E_x = C_{xs} C_{ss}^{-1}. \quad (20)$$

The matrix E_x estimates the turbulence profile x from the WFS measurement vector s , and the matrix F_x then fits a DM actuator command vector a to this estimated value. This decomposition of the minimum variance wavefront reconstructor has been commented upon previously [13], and it offers several useful insights into the reconstruction process. Note, for example, that the estimation matrix E_x is independent of the value of the weighting matrix W , and that the fitting matrix F_x likewise does not depend upon the statistics of the WFS measurement noise and atmospheric turbulence.

For relatively small order AO system, the matrices F_x and E_x can be computed explicitly to obtain the minimum variance reconstruction algorithm. For higher-order systems it is necessary to exploit the structure of the the matrices appearing in these definitions to obtain more efficient solutions, as described further below.

4 An efficient solution for $u = F_x v$

Eq. (19) for the matrix F_x is weakly similar to Eq. (3) for the least squares wavefront reconstructor, which suggests the possibility of applying sparse matrix techniques. The matrices H_x and H_a are sparse, since (as is suggested by Fig 2a), each element of the phase profile vector ϕ is a function of only a few elements of the atmospheric turbulence vector x and the DM actuator command vector a . However, the structure of the matrix F_x is complicated by the presence of the weighting matrix W and the extra term $N_w N_w^T$. The following two subsections develop the structure of W in greater detail, and summarize how sparse matrix methods may still be applied to find efficient solutions for $u = F_x v$.

A Weighting matrix structure

As described earlier, the matrix W is a positive semidefinite matrix chosen so that the quantity $\phi^T W \phi$ equals the mean-square piston-removed wavefront error for the phase profile(s) ϕ . In the wide field-of-view case, the vector ϕ will be a concatenation of phase profiles ϕ^1, \dots, ϕ^n from n different evaluation directions

$$\phi = \begin{pmatrix} \phi^1 \\ \vdots \\ \phi^n \end{pmatrix}, \quad (21)$$

where each profile ϕ^i is evaluated on the identical set of grid points in the telescope aperture plane. To account for aperture edge effects more precisely and obtain more accurate values for σ^2 with a limited number of grid points, each discrete phase profile ϕ^i may be associated with a continuous profile $\varphi^i(r)$ through the relationship

$$\varphi^i(r) = \sum_j \phi_j^i e_j(r), \quad (22)$$

where the terms $e_j(r)$ are localized ‘‘influence functions’’ associated with each point in the grid. The goal of this section is to determine a value for the weighting matrix W so that the value of $\phi^T W \phi$ is actually equal to a weighted sum of the mean-square, piston-removed values of the functions $\varphi^i(r)$, evaluated over the continuous telescope aperture.

The weighting matrix W is block diagonal, since there is no cross coupling between the distinct ϕ^i and ϕ^j in determining the mean-square phase error. Each block is identical up to an overall scale factor, since the telescope aperture function is identical for each ϕ^i but we may wish to assign different importance to each of the different phase profiles. The mean square phase error formula therefore takes the form

$$\phi^T W \phi = \sum_{i=1}^n w_i (\phi^i)^T V \phi^i, \quad (23)$$

where w_1, \dots, w_n are scalar weights and the matrix V is the same for each phase profile. The coefficients of V are defined so that the quantity $(\phi^i)^T V \phi^i$ is equal to the mean-square, piston removed value of $\varphi^i(r)$ averaged over the continuous telescope aperture. With the telescope aperture function denoted as $A(r)$, the matrix V must satisfy the relationship

$$(\phi^i)^T V \phi^i = \int dr A(r) \left[\varphi^i(r) - \int dr' A(r') \varphi^i(r') \right]^2, \quad (24)$$

where we have assumed for simplicity that the aperture function $A(r)$ has been normalized so that

$$\int dr A(r) = 1. \quad (25)$$

The coefficients of the matrix V may now be determined by substituting Eq. (22) into Eq. (24), expanding the square, and interchanging the order of summation and integration. The result is given by the formula

$$V = V_0 - V_1 V_1^T, \quad (26)$$

where the coefficients of the matrix V_0 and the column vector V_1 are defined by the integrals

$$(V_0)_{ij} = \int dr A(r) e_i(r) e_j(r), \quad (27)$$

$$(V_1)_i = \int dr A(r) e_i(r). \quad (28)$$

The matrix V_0 is sparse, since it is assumed that the influence functions $e_i(r)$ are localized. On account of Eq. (23), the value of the overall weighting matrix W is now given by the formulas

$$W = W_0 - W_1 W_1^T, \quad (29)$$

$$W_0 = \text{diag}(w_1 V_0, \dots, w_n V_0), \quad (30)$$

$$W_1 = \text{diag}(\sqrt{w_1} V_1, \dots, \sqrt{w_n} V_1). \quad (31)$$

(Note that the blocks appearing in the definition of the matrix W_1 are column vectors, not square matrices.) The key features of this representation are the facts that: (i) the matrix W_0 is sparse, and (ii) the matrix W_1 is a low rank matrix with a small number (n) of columns. This structure enables the efficient evaluation of $u = F_x v$, as described in the following subsection. [24]

B Solution summary

According to Eq. (19) for the matrix F_x , solving the system $u = F_x v$ for u can be accomplished by first computing $u' = H_a^T W H_x v$ and then solving $u = (H_a^T W H_a + N_w N_w^T + kI)^{-1} u'$. Both steps may be solved more efficiently by utilizing the facts that the matrices H_a , H_x , and W_0 are sparse, and that the matrices W_1 and N_w are low rank matrices with only a few columns.

The first of these two steps is conceptually simplest, since it does not involve a matrix inversion. To evaluate the value of u' defined by

$$u' = H_a^T W H_x v, \quad (32)$$

we first rewrite the matrix $H_a^T W H_x$ in the form

$$H_a^T W H_x = M - UV^T, \quad (33)$$

where by Eq. (29) the values of M , U , and V are defined by the expressions

$$M = H_a^T W_0 H_x, \quad (34)$$

$$U = H_a^T W_1, \quad (35)$$

$$V = H_x^T W_1. \quad (36)$$

The matrix M is sparse because each of the matrices H_a , W_0 , and H_x are sparse, and both U and V are low rank matrices with only a few columns because the same is true of the matrix W_1 . It follows that the vector u' can be computed efficiently using the expression

$$u' = M v - U(V^T v), \quad (37)$$

Where all of the matrix-vector operations involve sparse matrices or matrices with only a few rows or columns.

Solving the second step

$$u = (H_a^T W H_a + N_w N_w^T + kI)^{-1} u' \quad (38)$$

also depends upon a similar sparse-plus-low-rank representation

$$H_a^T W H_a + N_w N_w^T + kI = M - UV^T, \quad (39)$$

where the values of the matrices M , U , and V are now given by

$$M = H_a^T W_0 H_a + kI, \quad (40)$$

$$U = \begin{pmatrix} H_a^T W_1 & N_w \end{pmatrix}, \quad (41)$$

$$V = \begin{pmatrix} H_x^T W_1 & -N_w \end{pmatrix}. \quad (42)$$

The matrix M is sparse because H_a and W_0 are sparse, and the matrices U and V are once again low rank with only a small number of columns. Thanks to this representation we may now apply the matrix inversion lemma

$$(M \mp UV^T)^{-1} = M^{-1} \pm M^{-1}U(I \mp V^T M^{-1}U)^{-1}(M^{-1}V)^T, \quad (43)$$

which may be verified by multiplying both sides of the equation by M and $(M \mp UV^T)$ and then simplifying until an identity is obtained. Substituting Eq.'s (39) and (43) back into Eq. (38) now yields the result

$$u = M^{-1}u' + \left\{ (M^{-1}U) \left\{ (I - V^T M^{-1}U)^{-1} \left[(M^{-1}V)^T u' \right] \right\} \right\}. \quad (44)$$

The matrices $M^{-1}U$ and $M^{-1}V$ can be efficiently precomputed since M is sparse and the matrices U and V have only a small number of columns. The matrix $(I - V^T M^{-1}U)$ can then be computed and inverted efficiently because both V and $M^{-1}U$ have only a small number of columns. Eq. (44) can then be solved efficiently once these quantities are precomputed, since M is sparse and all operations following the addition symbol involve only matrices of low rank.

We have not formally evaluated the reduction in complexity that can be obtained by the above methods, but numerical results on actual computation times will be presented below in Section 6. The greatest savings are actually in precomputing the matrices, since the number of operations necessary to compute $H_a^T W H_x$ and $H_a^T W H_a$ will scale as the *cube* of the dimension of a when utilizing conventional matrix multiplies.

5 An efficient solution for $u = E_x v$

This part of the derivation begins by finding a more explicit representation for the phase estimation matrix E_x . Using Eq. (9), this matrix may be rewritten in the form

$$\begin{aligned} E_x &= C_{xs} C_{ss}^{-1} \\ &= C_{xx} G_x^T (G_x C_{xx}^T G_x^T + C_{nn})^{-1} \\ &= (G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1})^{-1} G_x^T C_{nn}^{-1}, \end{aligned} \quad (45)$$

where the last equality may be verified by multiplying both sides of the expression by $G_x C_{xx}^T G_x^T + C_{nn}$ on the right and by $G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1}$ on the left, and then further simplifying until an identity is obtained. Sparse matrix methods are not immediately applicable to this representation for several reasons. Most importantly, the phase covariance matrix C_{xx} is in general non-sparse, and in fact has infinite eigenvalues for the usual case of the Kolmogorov turbulence spectrum. An approximation to the *inverse* of C_{xx} must be chosen which is sparse and still fairly accurate. Secondly, the structure of the phase-to-WFS influence matrix G_x and the noise covariance matrix C_{nn} are complicated by the tilt uncertainty problem for laser guide stars, and by the full aperture tip/tilt NGS wavefront sensors that must be included in LGS AO and MCAO systems as a consequence of this effect. The following two subsections describe methods for coping with these complications, and the final subsection of this chapter then outlines how $u = E_x v$ can be efficiently solved for u .

A Approximating the regularizing term C_{xx}^{-1}

1 Eliminating the Kolmogorov singularity

For a pure Kolmogorov turbulence spectrum, the phase covariance matrix C_{xx} is not defined because the pure piston mode will have an infinite variance for each turbulence layer [25]. Even if we restrict attention to a von Karman spectrum with a large but finite outer scale, the matrix C_{xx}^{-1} will be

ill conditioned, with the pure piston mode of each turbulence layer approximating an eigenvector with a near-zero eigenvalue. The pure piston mode for each turbulence layer is also an eigenvector with a zero eigenvalue for the matrix $G_x^T C_{nn}^{-1} G_x$, since these modes lie in the null space of G_x for any existing or postulated wavefront sensing concept. The purpose of this subsection is to describe how to add an additional sparse regularizing term to the sum $G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1}$ so that this matrix, which appears in the definitions of E_x , will be better conditioned.

To derive this regularizing term, the matrices L and Z are first defined by the equations

$$L_{ij} = \begin{cases} 1 & \text{If phase point } i \text{ is located on screen } j, \\ 0 & \text{otherwise,} \end{cases} \quad (46)$$

$$Z_{ij} = \begin{cases} 1 & \text{If phase point } i \text{ is the origin of screen } j, \\ 0 & \text{otherwise.} \end{cases} \quad (47)$$

Each of these matrices has a number of rows equal to the total dimension of the phase error vector x , and a number of columns equal to the number of turbulence layers. $G_x L = 0$ since each column of the matrix L is a pure piston mode for one of the turbulence layers, and we note also that $Z^T L = I$. Consequently we may write

$$\begin{aligned} (G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1} + Z Z^T)(I - L Z^T) &= G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1} - C_{xx}^{-1} L Z^T \\ &\approx G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1}, \end{aligned} \quad (48)$$

where the approximation follows because the pure piston modes for each turbulence layer are approximately eigenvalues for C_{xx}^{-1} with near-zero eigenvalues. Rearranging this expression yields the relationship

$$(G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1} + Z Z^T)^{-1} \approx (I - L Z^T)(G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1})^{-1}. \quad (49)$$

The right-hand-side of this expression consists of the matrix inverse appearing in the expression for E_x on the last line of Eq. (45), followed by the term $I - L Z^T$. The action of this second operator is to adjust the overall piston term of each phase screen layer to obtain a value of zero at the origin, which will have no impact upon the piston-removed accuracy of the phase estimate. The left-hand-side of Eq. (49) may therefore be used instead of the term $(G_x^T C_{nn}^{-1} G_x + C_{xx}^{-1})^{-1}$ to obtain a slightly modified version of the phase estimator E_x with improved conditioning. The possible loss of performance due to the approximation made in Eq. (48) will be investigated via simulations in Section 6 below.

2 A sparse covariance matrix approximation

For the Kolmogorov turbulence spectrum the covariance matrix C_{xx} and its inverse will be non-sparse and of full rank, so an approximation of some sort must be made to proceed using the sparse techniques proposed by this paper. To justify our approximation, we consider the case of a single continuous turbulence layer of infinite extent. In this limit case, the bilinear functional defined by the matrix C_{xx}^{-1} may be approximated as

$$\begin{aligned} u^T C_{xx}^{-1} v &= u^T \langle x x^T \rangle^{-1} v \\ &= \iint dr dr' u(r) v(r') \left(\langle x(r) x^T(r') \rangle^{-1} \right) \\ &= \iint d\kappa d\kappa' \hat{u}(\kappa) \hat{v}(\kappa') \left(\langle \hat{x}(\kappa) \hat{x}^*(\kappa') \rangle^{-1} \right) \end{aligned}$$

$$\begin{aligned}
&\propto \int d\kappa \hat{u}(\kappa) \hat{v}^*(\kappa) \kappa^{11/3} \\
&\approx \int d\kappa \left[\kappa^2 \hat{u}(\kappa) \right] \left[\kappa^2 \hat{v}(\kappa) \right]^* \\
&\propto \int dr \nabla^2 u(r) \nabla^2 v(r),
\end{aligned} \tag{50}$$

where the spatial Fourier transform of a function f has been denoted as \hat{f} . The second line follows because we have assumed an infinite, continuous phase screen, the third by the Plancherel theorem, and the fourth by the definition of the Kolmogorov turbulence spectrum. The fifth line is based upon the approximation $11/3 \approx 4$, and the final line follows because $(df/dx) = 2\pi\kappa_x i f$. In other words, the quantity $u C_{xx}^{-1} v^T$ is approximately proportional to the inner product between the Laplacians, or curvatures, of u and v .

Since this heuristic argument is valid for all choices of u and v , it suggests that we can approximate C_{xx}^{-1} in the case of a single discrete turbulence layer as

$$C_{xx}^{-1} \approx C^T C, \tag{51}$$

where C is proportional to a discrete approximation of the Laplacian operator as illustrated in Fig. 3, with a constant of proportionality depending upon the strength of the turbulence. For a multilayer turbulence profile, the matrix C becomes a weighted sum of one such term per layer. The approximation to C_{xx}^{-1} is quite sparse, since the matrix C has no more than 5 nonzero elements per row. This computational simplification was first suggested and evaluated in [18] for the case of a conventional NGS AO system, but the motivation provided by Eq. (50) above and the extension to more than one atmospheric turbulence layer were both overlooked. We note that the boundary conditions on C illustrated in Fig. 3 were chosen somewhat arbitrarily, but based upon the numerical results presented in section 6 below there is no motivation for optimizing them further.

Combining this approximation with the discussion in the preceding subsection, the phase estimation matrix we propose to evaluate is given by the expression

$$E'_x = (G_x^T C_{nn}^{-1} G_x + C^T C + Z Z^T)^{-1} G_x^T C_{nn}^{-1}. \tag{52}$$

The remainder of this section describes how E'_x can be evaluated using sparse matrix techniques, and Section 6 assesses the loss in phase estimation accuracy incurred by making these approximations.

B Matrix structure for laser guide stars

As mentioned previously, the noise covariance matrix C_{nn} is no longer diagonal for the case of laser guide star (LGS) WFS measurements due to the effect of LGS position uncertainty. The exact position of a LGS guide star projected into the sky is variable due to both fundamental effects of atmospheric turbulence and practical error sources in the laser system, and (at least for now) there is no independent means of measuring the actual position with any accuracy. This motion of the guide star is indistinguishable from overall wavefront tilt to the wavefront sensor, resulting in an additional source of measurement noise that is fully correlated between all of the subapertures of a particular WFS. The matrices C_{nn} and $G_x^T C_{nn}^{-1} G_x$ are no longer sparse when this noise term is included, complicating the application of sparse methods to the evaluation of the phase estimation matrix E'_x .

More importantly from a practical perspective, one or more natural guide stars must always be included in the guide star constellation of a LGS AO system to measure the overall tip/tilt mode of the wavefront error, since the tip/tilt measurement from the LGS WFS is very noisy due to the uncertainty of the guide star position. The overall tip/tilt measurement from a NGS WFS with only

one or a few subapertures will depend upon phase values distributed across the entire telescope aperture, further degrading the sparseness of the term $G_x^T C_{nn}^{-1} G_x$ appearing in the formula for E'_x . Fortunately, both of these effects are low rank adjustments to the matrix, permitting the use of the matrix inversion lemma as used in Section 4 above. The remainder of this subsection introduces the notation required to discuss the natural- and laser guide star components of the WFS measurement vector s , and describes the form of the noise covariance matrix C_{nn}^{-1} when LGS position uncertainty is present.

For the most general case of either a conventional AO or MCAO system using a combination of natural and/or laser guide stars, it is still possible to decompose the WFS measurement vector s into a higher-order component s_h and a tip/tilt component s_t . The component s_h includes the measurements from all wavefront sensors with two or more subapertures, while s_t consists of the tip/tilt measurements from NGS sensors with only a single subaperture. The dimensionality of s_t is consequently equal to twice the number of tip/tilt natural guide stars, typically much smaller than the dimensionality of the higher-order component s_h . With this definition the WFS measurement model defined by Eq. (9) can be rewritten as

$$\begin{aligned} s &= \begin{pmatrix} s_h \\ s_t \end{pmatrix} \\ &= \begin{pmatrix} G_h \\ G_t \end{pmatrix} x + \begin{pmatrix} n_h \\ n_t \end{pmatrix}, \end{aligned} \quad (53)$$

where the dimensionality of n_t and the number of rows of G_t are also equal to twice the number of tip/tilt NGS. The matrix G_t is consequently of low rank relative to the overall size of the matrix G_x .

With the WFS measurement noise written as in Eq. (53), the form of the noise covariance matrix now becomes

$$C_{nn} = \begin{pmatrix} N_h + \sigma_t^2 T T^T & 0 \\ 0 & N_t \end{pmatrix}. \quad (54)$$

The terms N_h and N_t describe the statistics of the noise within the higher-order and tip/tilt wavefront sensors themselves and are diagonal matrices. The effect of LGS position uncertainty is captured in the term $\sigma_t^2 T T^T$, where σ_t is the RMS one-axis position uncertainty for each LGS, and the columns of the matrix T are the modes of WFS measurement noise induced by the LGS position errors. There are two such modes for each LGS, corresponding to the tip and tilt (or x - and y) position errors. Each column of T is $\{0, 1\}$ -valued, with the 1's matching those elements of s_h which are the x - or y - measurements for a particular laser guide star. The point for the current discussion is that the matrix T is of low rank, since the number of laser guide stars is very small compared to the total number of WFS measurements.

[The above representations of s and C_{nn} can also be applied to the case of a conventional AO or MCAO system with only higher order NGS WFS measurements, simply by taking $\sigma_t = 0$ and setting G_t and N_t to be empty matrices.]

The inverse of the noise covariance matrix C_{nn} will be needed to evaluate the phase estimation matrix E'_x . As before, the non-sparse term in this inverse may be evaluated using the matrix inversion lemma, yielding the result

$$\begin{aligned} (N_h + \sigma_t^2 T T^T)^{-1} &= N_h^{-1} - \sigma_t^2 N_h^{-1} T (I + \sigma_t^2 T^T N_h^{-1} T)^{-1} (N_h^{-1} T)^T \\ &\rightarrow N_h^{-1} - N_h^{-1} T (T^T N_h^{-1} T)^{-1} (N_h^{-1} T)^T \quad \text{as } \sigma_t^2 \rightarrow \infty \end{aligned} \quad (55)$$

The assumption that the RMS LGS position uncertainty is effectively infinite is only slightly conservative in practical applications; σ_t will be no smaller than the actual RMS value of the turbulence-

induced tip/tilt errors unless the size of the laser launch telescope is actually larger than the aperture of the AO system. The measured position of the LGS is consequently of little or no use in estimating the tip/tilt error, which is the motivation for including one or more natural guide stars as part of the guide star constellation in the first place. The computational advantage gained by treating σ_t as infinite is that the second line of Eq. (55) is a sparse matrix with a low rank adjustment, since N_h is diagonal and the matrix T has only a few columns.

C Solution summary

With the above preparations we are finally ready to describe the efficient evaluation of $u = E'_x v$ using sparse matrix methods. As in Section 4 for the actuator fitting operator F_x , the solution is given in two steps. The first step is the solution of the system

$$u' = G_x^T C_{nn}^{-1} v. \quad (56)$$

Using the representations of G_x and C_{nn}^{-1} developed in Eq.'s (53) through (55) above, the operator appearing in Eq. (56) may be rewritten in the form

$$G_x^T C_{nn}^{-1} = \begin{pmatrix} G_h^T N_h^{-1} (I - T P_T) & G_t^T N_t^{-1} \end{pmatrix}, \quad (57)$$

where the noise weighted LGS tilt projection operator P_T is defined as

$$P_T = (T^T N_h^{-1} T)^{-1} T^T N_h^{-1}. \quad (58)$$

Proceeding as in Section 4 for the actuator fitting problem, we now write the matrix $G_x^T C_{nn}^{-1}$ in the form

$$G_x^T C_{nn}^{-1} = M - UV^T, \quad (59)$$

where the terms M , U , and V are defined by the formulas

$$M = \begin{pmatrix} G_h^T N_h^{-1} & G_t^T N_t^{-1} \end{pmatrix}, \quad (60)$$

$$U = G_h^T N_h^{-1} T, \quad (61)$$

$$V = \begin{pmatrix} P_T^T \\ 0 \end{pmatrix} = \begin{pmatrix} N_h^{-1} T (T^T N_h^{-1} T)^{-1} \\ 0 \end{pmatrix}. \quad (62)$$

The matrix M is sparse because G_h is sparse, N_h is diagonal, and G_t is low rank (with only a few rows). The matrices U and V are low rank because the matrix T has only a few columns. This representation allows u' to be computed efficiently from v using Eq. (37) above.

[In passing, it may be interesting to note that the operator $(I - T P_T)$ appearing in Eq. (57) above has the effect of subtracting off the noise-weighted overall wavefront tilt from the LGS WFS measurements as the first step of the wavefront reconstruction algorithm. This preprocessing step, derived here using the matrix inversion lemma, may be considered ‘‘intuitively obvious,’’ and is already found in the reconstruction algorithms used for many LGS AO systems and simulations.]

The second and final step in evaluating $u = E'_x v$ is to efficiently solve the system

$$u = (G_x^T C_{nn}^{-1} G_x + C^T C + ZZ^T)^{-1} u'. \quad (63)$$

Using Eq.'s (53) and (57) above we may write

$$G_x^T C_{nn}^{-1} G_x + C^T C + ZZ^T = M - UV^T, \quad (64)$$

where the matrices M , U , and V are this time defined as

$$M = G_h^T N_h^{-1} G_h + C^T C + Z Z^T, \quad (65)$$

$$U = \begin{pmatrix} G_h^T N_h^{-1} T & -G_t^T N_t^{-1} \end{pmatrix}, \quad (66)$$

$$V = \begin{pmatrix} G_h^T P_T^T & G_t^T \end{pmatrix}. \quad (67)$$

The matrix M is sparse because all of the matrices appearing on the right hand side of Eq. (65) are sparse. The matrices U and V are low rank with only a few columns because the same is true of the matrices T , G_t^T , and P_T^T . We may therefore proceed as in Section 4 for the actuator fitting operator, applying the matrix inversion lemma and Eq. (44) to solve for u in terms of u' using operations that involve only low rank matrices and back substitutions through the sparse matrix M .

6 Sample numerical results

This section summarizes initial numerical wavefront reconstruction results obtained using the sparse computational methods developed in Sections 3 through 5 above. The cases considered have been chosen to address four particular questions:

- How do the approximations introduced in Section 5 with respect to atmospheric turbulence statistics degrade the accuracy of wavefront reconstruction?
- How large (or small) is the improvement in computational efficiency relative to the conventional matrix multiply reconstructor (CMMR)? We have concentrated on the time necessary to precompute the reconstructor, since this is generally the limiting factor in performing AO simulations.
- How large are the computer memory requirements to store the sparse reconstruction algorithm (SRA)? Memory requirements are a concern because the SRA computes an estimate of the full atmospheric turbulence profile x , which may have a significantly larger dimensionality than the DM actuator command vector a .
- What new results can be obtained about the performance of extreme AO and MCAO for extremely large telescopes, assuming that the SRA enables simulations of high-order AO systems that were not previously possible?

The results obtained regarding these questions are summarized in the following subsections. Subsection A outlines the cases considered and some of the details of the simulation code. Subsection B summarizes results for low order NGS AO systems which indicate that the loss in performance due to approximating atmospheric turbulence statistics is entirely negligible. Subsection C and D then present results for extreme NGS AO and MCAO systems. These results indicate that the reduction in computation times is very significant, computer memory requirements are generally acceptable, and the trends in ExAO and MCAO performance with increasing AO system order are very, very gradual.

A Cases Considered

Table 1 presents the atmospheric turbulence profile used for the simulations described in this section. This 6-layer profile is based upon thermosonde and generalized SCIDAR measurements taken at Cerro Pachon, Chile, the site of the Gemini-South telescope [26]. The profile has been scaled to

obtain an r_0 of 16 cm at a wavelength of $0.50 \mu\text{m}$, corresponding roughly to median conditions at Cerro Pachon. The resulting value of θ_0 is 2.65 arc sec or $12.85 \mu\text{rad}$.

Table 2 summarizes the sample AO systems that have been evaluated. These include conventional NGS AO systems with telescope aperture diameters of 4, 8, 16, and 32 meters, and NGS and LGS MCAO systems with aperture diameters of 8 and 16 meters. The WFS subaperture width and DM actuator pitch for the conventional NGS AO systems was 0.5 meters, with performance evaluated for a single evaluation direction and an on-axis guide star. For the MCAO system performance was evaluated at the center, edges, and corners of a one square arc minute field of view, with the weights w_i assigned to the nine evaluation directions in Eq. (23) determined using Simpson’s rule. Five higher order guidestars were located at the center and corner of the 1 arc minute field for both the NGS and LGS MCAO systems. For LGS MCAO, these guide stars were placed at a range of 90 km, and four tip/tilt natural guide stars were also located at the edges of the 1 arc minute field. The subaperture width was also equal to 0.5 meters for the higher-order wavefront sensors in both the MCAO systems. Both MCAO systems also included two deformable mirrors at conjugate ranges of 0.0 and 5.16 km with actuator pitches of 0.5 meters, and a third DM at a range of 10.31 km with a 1.0 meter actuator pitch.

The code written to evaluate the sparse reconstruction algorithm (SRA) for these AO system consisted of three basic parts. The first part computed the influence matrices H_x , H_a , and G_x , the phase error weighting matrix W , and the approximate regularization matrix $C^T C$ for the given turbulence profile, aperture geometry, DM configuration, and guide star constellation. As illustrated in Fig. 2, the first four matrices were computed assuming linear spline influence functions. The matrix elements corresponding to boundary actuators, phase points, and subapertures were computed with the circular telescope aperture taken into account. The exact turbulence covariance matrix C_{xx} was also computed (with the full aperture piston mode removed to eliminate the singularity) for simulation cases with an aperture diameter of 8 meters or less.

The second part of the code then computed the SRA, as well as the CMMR for the 4- and 8-meter simulations. The final step was to evaluate the SRA, and for small cases also the CMMR, via Monte Carlo simulation, using 100 turbulence profile realizations generated as the Fourier transforms of white noise filtered by the Kolmogorov spectrum. Identical profiles were used to evaluate the SRA and the CMMR for the 4- and 8-meter cases.

The atmospheric turbulence profile sampling used to compute the reconstructors was at twice the spatial resolution of the deformable mirror actuators, *i.e.* grid point spacings of either 0.25 or 0.5 meters for the six turbulence layers. The simulations themselves used grids with spatial resolutions sixteen times finer than the DM actuator spacings, *i.e.* 0.03125 or 0.0625 meters. A few preliminary simulation results indicated that further increasing the resolution used to compute the reconstructors had a negligible effect on estimation accuracy, and that the “simulation fitting error” arising from unsampled high spatial frequency turbulence scaled as the two-thirds power of the grid spacing used in the simulation. We believe that this “simulation fitting error” results from omitting the spatial aliasing of the highest spatial frequency turbulence into the WFS measurements, which is a source of measurement error. Using this scaling law, the magnitude of this error with the chosen simulation grid spacings is estimated to be a mean-square OPD of about $0.0028 \mu\text{m}^2$, which is small compared with the effects studied in the following simulations.

B Results for Low order NGS AO

For an initial performance comparison between the SRA and the CMMR, the first AO system studied was a conventional, narrow field-of-view NGS AO system of order 8×8 with an on-axis guide star. The sources of wavefront error in these simulations were consequently WFS measurement noise and and DM/WFS fitting error, and we considered a range of WFS noise levels such that

the RMS wavefront estimation error due to measurement noise varied from about one-half to three times the RMS fitting error. The results of these simulations are presented in Table 3. Note that four significant digits are necessary to detect the performance variations between the SRA and the CMMR at low noise levels, and at the largest noise level considered the values of σ^2 for the two approaches still agree to within about 1%. At least in this case, the degradation in wavefront estimation accuracy due to the approximations made in deriving the SRA will be negligible for practical applications.

C Results for Extreme NGS AO

We next varied the order of the NGS AO system for a fixed, small value of WFS measurement noise. The results obtained are summarized in Table 4. For an order 16×16 system it is still possible to compute the CMMR without significant difficulty, and in this case the performance difference between the SRA and CMMR remains very, very small. For the SRA the value of σ^2 grows by about 8 per cent as the order of the NGS AO system increases from 16×16 to 64×64 . This trend indicates that the algorithm remains numerically stable. It is also qualitatively consistent with the expected slow growth in reconstructor noise gain as a function of AO system order, although we have not compared these values explicitly against standard scaling laws [12].

We note that the time required to compute the SRA increases somewhat more slowly than the three-halves power of the number of DM actuators, and that the time needed to compute the reconstructor for an order 64×64 NGS AO system is about 9 seconds on a 1 GHz Pentium III using Matlab 6 [27]. Significantly larger cases could easily be evaluated for truly extreme AO simulations. We have not recorded the time necessary to apply the reconstructor, but this will be essentially proportional to the number of nonzero coefficients in the sparse representations of the estimation and fitting matrices E_x and F_x , or equivalently the amount of memory needed to store these matrices. These storage requirements increased by a factor of about 26 while the number of DM actuators increased by a factor of $3461/257 = 13.5$. For the CMMR, the corresponding increase would be about $13.5^2 = 181$.

D Results for NGS and LGS MCAO

Finally, Table 5 summarizes the results obtained on wavefront estimation error and computational efficiency for the NGS and LGS MCAO systems with parameters as described in Table 2. We have considered systems of aperture diameters 8 and 16 meters with a fixed WFS subaperture size of 0.5 meters, yielding AO system orders of 16×16 and 32×32 respectively. For the order 16×16 case it is possible to compute and simulate both the CMMR and SRA methods, and the mean-square estimation errors for the two techniques are again virtually identical. We note that the simulated performance results for the SRA are actually very slightly superior to the CMMR in the case of LGS MCAO, presumably due to the effects of finite numerical precision in evaluating the reconstructors and the finite number of simulation trials used to obtain the performance estimates. Also, the performance of the SRA is almost independent of telescope aperture diameter for both NGS AO and LGS MCAO, suggesting that results for a 32 meter, order 64×64 system may very likely be similar as well. We are unable to explicitly evaluate the 32 meter case at this time due to computer memory limitations.

The computer memory requirements for the MCAO cases are very significantly greater than for the Extreme AO case considered in Table 4 because (i) the entire 3-dimensional turbulence profile must be evaluated for the MCAO case and (ii) the DM-to-WFS influence matrix G_x and the matrices derived from it are much less sparse thanks to the presence of multiple guide stars distributed over an extended field-of-view. The memory requirements predicted for the 32 meter case still appear to be feasible for existing computer systems, but not for our available PC's.

The increased fill factor for these matrices has negative implications for the effort required to compute the SRA as well, but the time required for the 16 meter, order 32×32 system with 2417 DM actuators and 4284 total WFS subapertures is still on the order of 1.7 to 1.9 hours, representing a factor of 8 improvement over the CMMR. Extrapolating these results via power laws to an order 64×64 MCAO system yields a predicted improvement factor of about 30 or 40 to 1. For the clock speed of our Pentium III, this would reduce the time needed to compute the reconstructor from about one month to the order of one day.

Acknowledgments

This research was supported by the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the National Science Foundation (NSF) on behalf of the Gemini Partnership: the United States National Science Foundation, the United Kingdom Particle Physics and Astronomy Research Council (PPARC), the Canadian National Research Council (NRC), the Chilean Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), the Australian Research Council (ARC), the Argentinean Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), and the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

1. J. W. Hardy, J. E. Lefevbre, and C. L. Koliopoulous, "Real time atmospheric turbulence compensation," *J. Opt. Soc. Am.* **67**, 360-369 (1977).
2. G. Rousset, J.-C. Fontanella, P. Kern, P. Gigan, F. Rigaut, P. Lena, C. Boyer, P. Jagourel, J.-P. Gaffard, and F. Merkle, "First diffraction-limited astronomical images with adaptive optics," *Astron. Astrophys.* **230**, 29-32 (1990).
3. M. C. Roggemann and B. Welsh, *Imaging through turbulence* (CRC Press, 1996).
4. D. L. Fried, "Least-squares fitting a wave-front distortion estimate to an array of phase difference measurements," *J. Opt. Soc. Am.* **67**, 370-375 (1977).
5. R. Hudgin, "Wave-front reconstruction for compensated imaging," *J. Opt. Soc. Am.* **67**, 375-378 (1977).
6. J. Hermann, "Least-squares wave-front errors of minimum norm," *J. Opt. Soc. Am.* **70**, 28-35 (1980).
7. E. P. Wallner, "Optimal wave-front correction using slope measurement," *J. Opt. Soc. Am.* **73**, 1771-1776 (1983).
8. B. M. Walsh and C. S. Gardner, "Effects of turbulence-induced anisoplanatism on the imaging performance of adaptive-astronomical telescopes using laser guide stars," *J. Opt. Soc. Am. A* **8**, 69-80 (1991).
9. G. A. Tyler, "Merging: a new method for tomography through random media," *J. Opt. Soc. Am. A* **1**, 409-424 (1994).
10. J. M. Beckers, "Increasing the size of the isoplanatic patch with multi-conjugate adaptive optics," in *Proceedings of European Southern Observatory Conference and Workshop on Very Large Telescopes and Their Instrumentation*, M.-H. Ulrich, ed., Vol. 30 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 1988), pp. 693-703.
11. D. C. Johnston and B. M. Welsh, "Analysis of multi-conjugate adaptive optics," *J. Opt. Soc. Am. A* **11**, 394-408 (1994).
12. B. L. Ellerbroek, "First order performance evaluation of adaptive-optics systems for atmospheric turbulence compensation in extended field-of-view astronomical telescopes," *J. Opt. Soc. Am. A* **11**, 783-805 (1994).
13. T. Fusco, J.-M. Conan, G. Rousset, L. M. Mugnier, and V. Michau, "Optimal wave-front reconstruction strategies for multi-conjugate adaptive optics," *J. Opt. Soc. Am. A* **18**, 2527-2538 (2001).

14. B. L. Ellerbroek, "Methods for correcting tilt anisoplanatism in laser-guide-star-based multi-conjugate adaptive optics," *J. Opt. Soc. Am. A* **18**, 2539-2547 (2001).
15. G.M. Cochran, "Sparse Matrix Techniques in Wavefront Reconstruction," Report No. TR-668, the Optical Sciences Company (Anaheim, CA, 1986).
16. L. Schmutz, B. M. Levine, A. Wirth, and C. Strandley, "Adaptive optics concepts for extremely large aperture telescopes," in *Backskog Workshop on Extremely Large Telescopes*, T. Andersen, A. Ardeberg, and R. Gilmozzi, eds., Vol. 57 of ESO Conference and Workshop Proceedings, 217-223 (European Southern Observatory, Garching, Germany, 1999).
17. L. Gilles, C. R. Vogel, and B. L. Ellerbroek, "Iterative Algorithms for Large Scale Wave Front Reconstruction," in *OSA Trends in Optics and Photonics (TOPS) Vol. 67, Signal Recovery and Synthesis*, OSA Technical Digest, Postconference Edition (Optical Society of America, Washington DC, 2001), pp. 100-101.
18. B. L. Ellerbroek, "Comparison of Least Squares and Minimum Variance Wavefront Reconstruction for Atmospheric Turbulence Compensation in the Presence of Noise," Report No. TR721R, the Optical Sciences Company (Anahiem, CA, 1986).
19. Gregory M. Cochran, "Sparse Matrix Techniques Applied to Deconvolution," *Computers and Electrical Engineering* **18**, 499-505 (1992).
20. R. Flicker, F.J. Rigaut, and B.L. Ellerbroek, "Comparison of multi-conjugate adaptive optics configurations and control algorithms for the Gemini South 8-m telescope," in *Adaptive Optical Systems and Technology*, Peter L. Wizinowich, ed., Proc. SPIE 4007, 4007-1032 (2000).
21. Gene H. Golub and Charles F. van Loan, *Matrix Computations* (John Hopkins Univ. Press, Baltimore, MD, 1996), pp. 18.
22. A. George and J. Liu, *Computer Solutions of Large Symmetric Positive Definite Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1981).
23. S. Pissantesky, *Sparse Matrix Technology* (Academic Press, Orlando, FL, 1984), Chap. 4.
24. The weighting matrix W derived in this section includes the global tip and tilt modes in the calculation of the mean-square wavefront error. This is appropriate for AO applications to long-exposure imaging, where random, time-varying tip and tilt errors will degrade image quality in the same fashion as higher-order wavefront errors. For short exposure imaging it may be more desirable to consider tip/tilt and higher order wavefront aberrations separately. In this case the weighting matrix W for the mean-square, higher-order wavefront error will still take the form defined by Eq.'s (29) through (31), except that the matrix V_1 now has three columns instead of one. The computational methods developed in the remainder of this paper are consequently still applicable.

25. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66**, 207-211 (1976).
26. J. Vernin, A. Agabi, R. Avila, M. Azouit, R. Conan, F. Martin, E. Masciadri, L. Sanchez, and A. Ziad, "1998 Gemini site testing campaign: Cerro Pachon and Cerro Tololo," Gemini Document RTP-AO-G0094 (Gemini Observatory, Hilo, Hawaii, 2000).
27. Matlab 6 no longer supports the "FLOPS" command included in previous versions to report the exact number of floating point operations required by an algorithm. No other computationally significant tasks were running on this dual processor system during the timing tests reported in this paper.

Table 1. Atmospheric turbulence profile used for simulations. This table lists the altitudes and relative weights for a six-layer atmospheric turbulence profile derived from thermosonde and generalized SCIDAR measurements at Cerro Pachon, Chile. The overall profile was scaled to yield $r_0 = 16$ cm at $\lambda = 0.5 \mu\text{m}$. The corresponding isoplanatic angle is $\theta_0 = 12.85 \mu\text{rad}$.

Layer	Altitude, km	Relative layer weight
1	0.00	0.652
2	2.58	0.172
3	5.16	0.055
4	7.73	0.025
5	12.89	0.074
6	15.46	0.022

Table 2. Simulated AO system parameters. See the text for further details.

System	NGS AO	MCAO
Evaluation field	On-axis	60 arc second square
Aperture diameter, m	4, 8, 16, 24, 32	8, 16
DM conjugate ranges, km	0.0	0.0, 5.16, 10.31
DM interactuator spacing, m	0.5	0.5, 0.5, 1.0
Higher order guidestars:		
Number	1	5
Subaperture spacing, m	0.5	0.5
Directions, arc min	(0, 0)	(0, 0) and ($\pm 0.5, \pm 0.5$)
Range, km	∞	∞ (NGS) or 90 (LGS)
Tip/tilt guidestars:		
Number	0	0 (NGS) or 4 (LGS)
Directions, arc min	–	($\pm 0.5, 0$) and ($0, \pm 0.5$)
Range, km	–	∞

Table 3. Reconstructor performance vs. WFS measurement noise for an order 8×8 NGS AO system. This table compares the residual mean square wavefront error σ^2 due to the combined effects of fitting error and WFS measurement noise for the conventional matrix multiply implementation of the minimum variance estimator (CMMR) and the sparse reconstruction algorithm (SRA) described in this paper. The approximations made in modeling atmospheric turbulence statistics for the latter algorithm have only a very modest effect on the mean-square wavefront estimation error σ^2 .

WFS noise, arc sec	CMMR $\sigma^2, \mu\text{m}^2$	SRA $\sigma^2, \mu\text{m}^2$
0.02	0.01304	0.01306
0.04	0.01734	0.01737
0.08	0.03218	0.03229
0.16	0.07834	0.07941

Table 4. SRA wavefront fitting error vs. order of correction for a conventional NGS AO system. This table summarizes the performance of the sparse reconstruction algorithm (SRA) for a conventional NGS AO scenario where the only significant source of wavefront error is the finite spatial resolution of the DM actuators and WFS subapertures. The natural guide star is coincident with the evaluation direction, and the noise equivalent angle for the WFS is an almost negligible 0.02 arc seconds. The DM actuator spacing is held constant, so the AO order of correction is proportional to the telescope aperture diameter. The SRA estimation error is virtually identical with the conventional matrix multiply minimum variance reconstructor (CMMR) for the case of the order 16×16 AO system, and increases fairly gradually with increasing telescope aperture diameter. The computation times and memory requirements for the SRA grow much less rapidly than the $O(n^3)$ and $O(n^2)$ scaling laws that apply for the case of the CMMR.

Aperture diameter, m	Order of correction	DM actuators	CMMR $\sigma^2, \mu\text{m}$	SRA $\sigma^2, \mu\text{m}$	SRA computation time, sec	SRA memory requirements, MB
8	16×16	257	0.01334	0.01336	0.25	0.72
16	32×32	921	–	0.01392	1.17	3.81
24	48×48	1981	–	0.01416	3.60	NA
32	64×64	3461	–	0.01440	8.94	19.30

Table 5. Results and scaling law predictions for CMMR and SRA performance for MCAO systems. The simulated MCAO parameters are summarized in Table 2 above. The parenthesized values are extrapolations based upon the standard $O(n^2)$ and $O(n^3)$ power laws for the CMMR, and 2-point power laws curve fits for the SRA.

	NGS MCAO			LGS MCAO		
Aperture diameter, m	8	16	32	8	16	32
System order	16×16	32×32	64×64	16×16	32×32	64×64
Total DM actuators	789	2417	(9700)	789	2417	(9700)
Total WFS subapertures	1020	4280	(17120)	1020	4280	(17120)
CMMR σ^2 , μm	0.01904	–	–	0.02220	–	–
Time to compute CMMR, hours	0.51	(14.66)	(947.66)	0.51	(14.66)	(947.66)
Memory to store CMMR, MB	14	(129)	(2071)	13.8	(130)	(2086)
SRA σ^2 , μm	0.01932	0.01877	–	0.02190	0.02220	–
Time to compute SRA, hours	0.19	1.92	(33.90)	0.21	1.68	(22.20)
Memory to store SRA, MB	112	680	(6380)	99	561	(4852)

Fig. 1. Computation requirements for conventional and sparse calculations of classical least squares reconstruction algorithms. These results are for a square aperture geometry and the so-called “Hudgins” or “shearing interferometer” wavefront sensor geometry. The number of floating point operations needed to compute the control algorithm using a conventional matrix inversion has been evaluated for systems of order 100, 225, and 400 and extrapolated using the predicted third-order power law. The number of operations necessary for the sparse matrix factorization has been explicitly computed for AO systems of up to order 90,000, and scales with the three-halves power of the order of the system.

Fig. 2. Influence matrix models. Part (a) of the figure illustrates the relationship between the turbulence phase screen vector x , the DM actuator command vector a , and the residual phase error vector ϕ . These three vectors are defined as values on grids of points in the planes of the phase screens, the DM conjugate locations, and the telescope aperture, respectively. The influence matrices H_x and H_a are defined by tracing rays through the phase screens and mirrors as illustrated. Part (b) of the figure illustrates the similar relationship between the phase screen vector x and the WFS measurement vector s . In this case rays are traced from the guide star(s) through the phase screen(s) to obtain a wavefront in the telescope aperture plane, and the WFS measurements are then computed as the average x - and y wavefront gradients over each subaperture.

Fig. 3. Discrete Laplacian operator. This figure illustrates the coefficients for two rows of the discrete Laplacian, or curvature, operator C appearing in Eq. (51). The dots represent the grid points of the discrete turbulence layer. The values printed in a regular font are the 5 nonzero coefficients to compute the curvature of the phase profile at the interior grid point A . The italicized values are the nonzero coefficients to compute the curvature at the grid point B on the boundary of the phase profile. The coefficients which should be assigned to grid points laying outside of the boundary have been “folded over” back into the grid so that the sum of the coefficients remains zero.

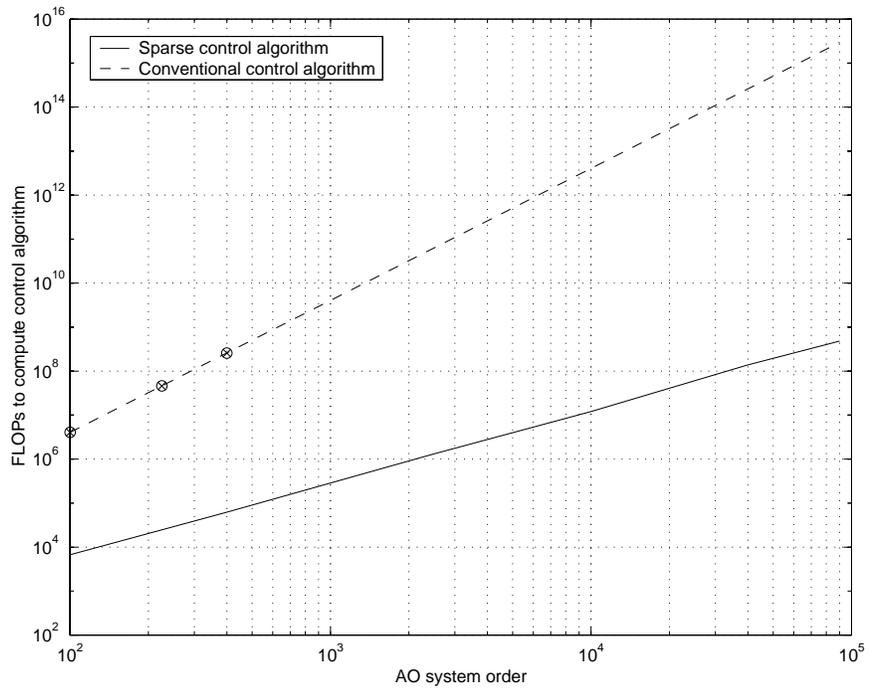


Figure 1

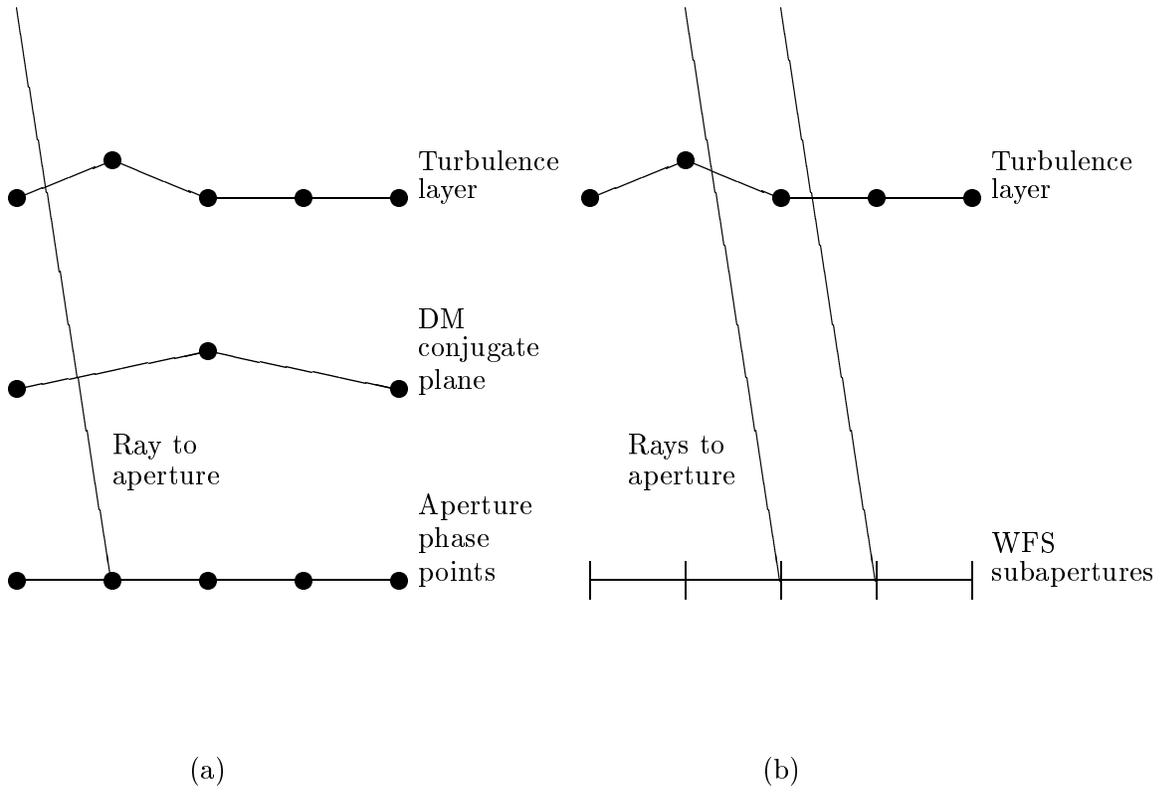


Figure 2

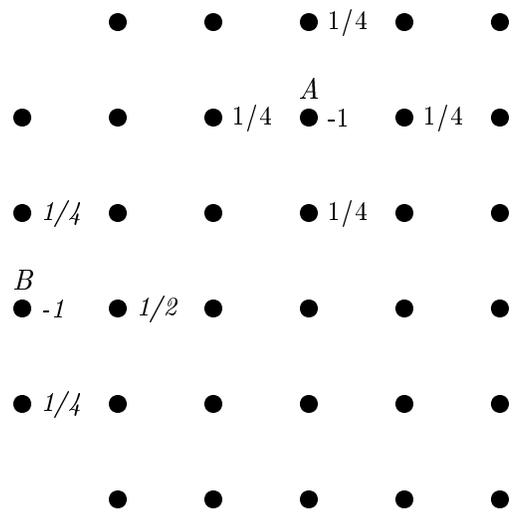


Figure 3