

Layer-Oriented Multigrid Wavefront Reconstruction Algorithms for Multi-Conjugate Adaptive Optics

L.Gilles^a, B.L.Ellerbroek^b and C.R.Vogel^c

^aMichigan Technological University, Dept. of Electrical and Computer Engineering,
Houghton, MI 49931-1295

^bGemini Observatory, 670 N. A'Ohoku Place, Hilo, HI 96720

^cMontana State University, Dept. of Mathematical Sciences,
Bozeman, MT 59717-2400

ABSTRACT

Multi-conjugate adaptive optics (MCAO) systems with $10^4 - 10^5$ degrees of freedom have been proposed for future giant telescopes. Using standard matrix methods to compute, optimize, and implement wavefront control algorithms for these systems is impractical, since the number of calculations required to compute (apply) the reconstruction matrix scales with the cube (square) of the number of AO degrees of freedom. In this paper, we develop an iterative sparse matrix implementation of minimum variance wavefront reconstruction for telescope diameters up to 32m with more than 10^4 actuators. The basic approach is the preconditioned conjugate gradient method, using a multigrid preconditioner incorporating a layer-oriented (block) symmetric Gauss-Seidel iterative smoothing operator. We present open-loop numerical simulation results and analyze algorithm convergence.

Keywords: Multi-Conjugate Adaptive Optics, Wavefront Reconstruction, Multigrid

1. INTRODUCTION

Conventional Adaptive Optics (AO) has now successfully been implemented on ground-based imaging systems, and there is a growing interest in developing, what will be the next step in imaging through volume turbulence, Multi-Conjugate Adaptive Optics (MCAO). Using multiple guide stars and multiple deformable mirrors (DM's), MCAO aims at breaking the small field of view (FoV) barrier inherent in conventional AO and the conical anisoplanatism inherent when using laser guide stars (LGS's). Such systems represent a huge leap in complexity from single mirror systems, and the task of creating an optimal scheme to control them is far from being trivial.

Wavefront reconstructors estimate a set of DM commands from wavefront sensor (WFS) slope measurements. Real-time wavefront reconstructors employ either a vector-matrix-multiply (VMM) or an iterative algorithm to compute the commands. Direct evaluation of the control matrix becomes increasingly difficult as the number of actuators n increases, since the processing time to compute the matrix and apply it to sensor data with a VMM scales approximately respectively as n^3 and n^2 (i.e. D^6 and D^4 where D is the aperture diameter) when standard matrix multiplication and inversion methods are employed. The control matrix may be computed in a few hours to a few days on a fast workstation for an AO system with $n = 10^3$, but MCAO systems on extremely large telescopes (ELT's) may have $n = 10^4 - 10^5$. Relying upon improvements in processor speed and parallel computing to cope with a factor of $10^3 - 10^6$ increase in computation requirements is clearly wasteful, and the development of new iterative algorithms and advanced numerical and mathematical techniques is highly desirable to help develop and evaluate conceptual designs for MCAO on ELT's.

For conventional AO systems with a single DM and WFS, significant improvements in computational efficiency are possible by exploiting the sparse structure of the deformable mirror influence matrices and atmospheric turbulence inverse covariance matrices appearing in these calculations. Sparse matrix techniques exploiting the sparse structure of the deformable mirror influence matrices, have reduced the overall cost of evaluating and applying the classical least-squares wavefront control matrix defined as the pseudoinverse of the

Send correspondence to L.Gilles

DM-to-WFS influence matrix, to $\mathcal{O}(n^{3/2})$. For MCAO systems employing multiple DM's and WFS's, the cost of computing and applying the least-squares reconstructor is a much larger constant multiple of $n^{3/2}$, and the reconstructor suffers from WFS noise amplification for weakly sensed modes. Some sort of filtering of the pseudoinverse like a singular value decomposition (SVD) appears necessary, but this cancels out the computational advantages of the sparse techniques. Optimum results can be achieved using a minimum variance estimator that includes a regularization term based upon the second-order statistics of atmospheric turbulence and WFS measurement noise.

Sparse matrix techniques have successfully been applied to the MCAO minimum variance reconstructor, and have reduced the overall computations to $\mathcal{O}(n^{3/2})$ using a small adjustment of the turbulence statistics so that the regularization term appearing in the optimal estimator is sparse.¹ For extreme adaptive optics (ExAO) systems (ultra-high-contrast AO system using a single high-order DM and WFS), a minimum variance reconstructor based on an iterative multigrid (multiresolution) method scaling as $n \log n$ has also been shown to be very effective,² and a Fourier transform reconstructor scaling as $n \log n$ has also been proposed.³ This last method directly inverts the phase-to-WFS influence matrix in the Fourier domain using an extension method, but suffers from noise amplification.

In this paper, we address the mathematical and computational issues for MCAO systems (in particular the dramatic amount of fill-in occurring in the Cholesky factorizations¹ due to the long-range cross-coupling between layers), and report on new ideas for fast and efficient iterative algorithmic approaches based on combined sparse matrix and multigrid techniques that scale in overall computational complexity and real time implementation as $\mathcal{O}(n \log n)$, or even $\mathcal{O}(n)$ with small constant multiply if the sparse approximation to the turbulence statistics made in Ref.1 is adopted. The basic approach is the conjugate gradient method with a multigrid preconditioner to speed up convergence, and a layer-oriented (block) symmetric Gauss-Seidel smoother inspired by the recent layer-oriented control algorithms for MCAO.

2. PROBLEM STATEMENT

A schematic MCAO system is illustrated in Fig.1. In this system, the WFS's are optically conjugate to the telescope's primary mirror and the DM's to different altitudes z_m^a in the atmosphere. The WFS's collect light from several LGS's located within the FoV at an altitude z_g . Turbulent layers are located at distinct ranges z_l^o .

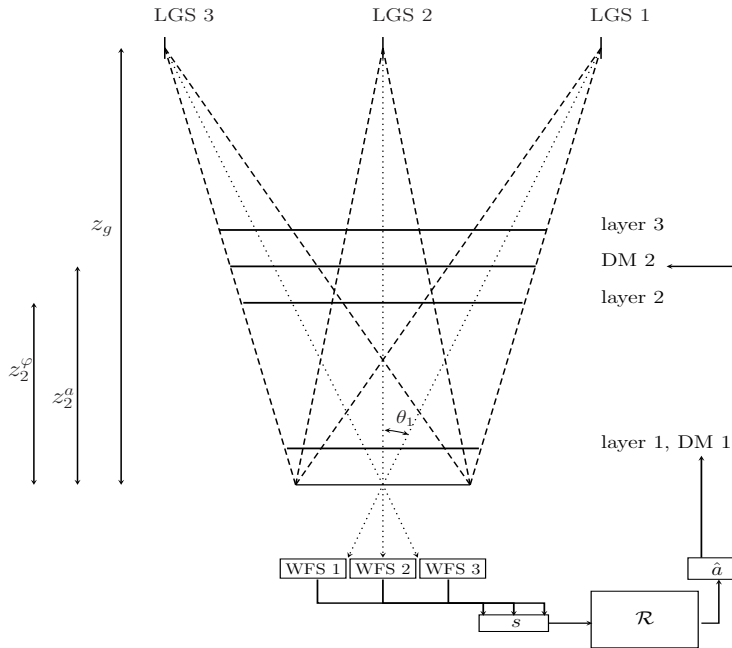


Figure 1. Schematic MCAO system.

Minimum variance wavefront reconstructors are now favored for MCAO.⁴ For the most general case of either a natural guide star (NGS) or a laser guide star (LGS) MCAO system with M DM's, H high-order WFS's, T full-aperture tip/tilt WFS's (apply only to the LGS case) and L atmospheric layers, with $N \geq L \geq M$, where $N = H + T$ is the total number of WFS's, these reconstructors are block matrix systems of the form

$$\hat{a} = \mathcal{R} \begin{pmatrix} s \\ \tilde{s} \end{pmatrix}, \quad \mathcal{R} = (\mathcal{R}_{mn})_{\substack{1 \leq m \leq M \\ 1 \leq n \leq N}}, \quad (2.1)$$

where \hat{a} , s and \tilde{s} are concatenations of respectively M, H and T subvectors corresponding to the estimated actuator commands a_m ($1 \leq m \leq M$), high-order s_h ($1 \leq h \leq H$) and full-aperture tip/tilt \tilde{s}_t ($1 \leq t \leq T$) WFS measurements. The beam print at layer l in direction θ_n together with the metapupil are shown in Fig.2.

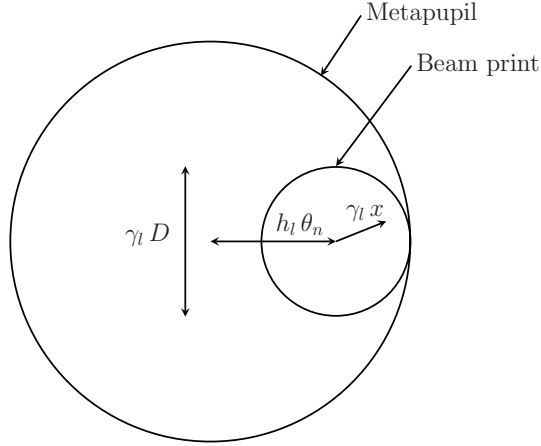


Figure 2. Metapupil and beam print. x is the pupil-plane coordinate, $\gamma_l = 1 - h_l/z_g$ where h_l and z_g are respectively layer l and LGS altitudes.

For open-loop wavefront sensing, the measurement vector is modeled as

$$\begin{pmatrix} s \\ \tilde{s} \end{pmatrix} = \mathcal{G} \varphi + \begin{pmatrix} \eta \\ \tilde{\eta} \end{pmatrix}, \quad \mathcal{G} \stackrel{\text{def}}{=} \begin{pmatrix} (I - P)G \\ \tilde{G} \end{pmatrix}, \quad (2.2)$$

where φ , η and $\tilde{\eta}$ are concatenations of respectively L, H and T subvectors corresponding to the true turbulence profile φ_l ($1 \leq l \leq L$), the high-order η_h ($1 \leq h \leq H$) and tip-tilt η_t ($1 \leq t \leq T$) WFS measurement noise, assumed of zero mean and finite second-order statistics. P is a block diagonal orthogonal tip-tilt projector with respect to the high-order WFS's inverse covariance matrix $C^{-1} = \langle \eta \eta^T \rangle^{-1}$ (applies only to the LGS case):

$$\begin{aligned} P &= \text{Diag}(P_h)_{1 \leq h \leq H}, \\ P_h &= K_h (K_h^T C_h^{-1} K_h)^{-1} K_h^T C_h^{-1}, \\ P_h^T &= C_h^{-1} P_h C_h \quad (\text{orthogonality}), \end{aligned} \quad (2.3)$$

where $C_h = \langle \eta_h \eta_h^T \rangle$ is the diagonal noise covariance matrix of high-order WFS h , and K_h is a rank 2 constant matrix (corresponding to the tip and tilt modes) whose number of rows is equal to the number of sensor h grid points. $(K_h^T C_h^{-1} K_h)$ is thus a 2×2 matrix and P is a low rank matrix ($\text{rank}(P) = 2H$) compared to the total number of high-order sensor grid points (number of rows of P). The phase-to-WFS influence matrices G and \tilde{G} have the block matrix representation:

$$G = (\Gamma_h H_{hl}^\varphi)_{\substack{1 \leq h \leq H \\ 1 \leq l \leq L}}, \quad \tilde{G} = \left(\tilde{\Gamma}_t \tilde{H}_{tl}^\varphi \right)_{\substack{1 \leq t \leq T \\ 1 \leq l \leq L}}, \quad (2.4)$$

where Γ_h is a second-order finite difference (FD) approximation of the gradient operator whose stencil depends on the high-order WFS's geometry, and $\tilde{\Gamma}_t$ is the full-aperture tip-tilt projector for NGS WFS t ($\text{rank}(\tilde{\Gamma}_t) = 2$). For NGS MCAO systems, P and \tilde{G} are null matrices, and \tilde{s} , $\tilde{\eta}$ null vectors. To compute phase values at arbitrary coordinates intersecting rays from direction θ_n , i.e. at $(\gamma_l x + \theta_n^x h_l, \gamma_l y + \theta_n^y h_l)$ for layer l , we use the interpolation formula

$$u^I(x, y) = \sum_i e_i(x, y) u_i, \quad (2.5)$$

where $e_i(x, y)$ is the influence function associated with grid point i . We are using here bilinear splines.⁵ A cross-section of bilinear spline influence functions is illustrated in Fig.3. Using a column or row grid point ordering scheme, equation (2.5) can be rewritten in matrix form

$$u^I = H u. \quad (2.6)$$

where u^I indicates interpolated quantity and the columns of the interpolation matrix H are the influence functions $e_i(x, y)$. In our context, we denote by

$$\varphi_{nl}^I = H_{nl}^\varphi \varphi_l, \quad (2.7)$$

the interpolated phase vector at layer l in direction θ_n . For laser guide stars at altitude z_g , $\gamma_l = (1 - h_l/z_g)$ (beam reduction if $z_g < \infty$). Note also that $H_{n1}^\varphi = I$ since layer 1 is at ground level and therefore no shift or beam reduction is needed. Figure 4 is a schematic illustrating a typical geometry using influence functions.

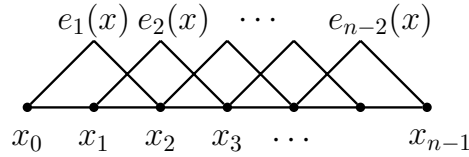


Figure 3. Bilinear spline influence functions.

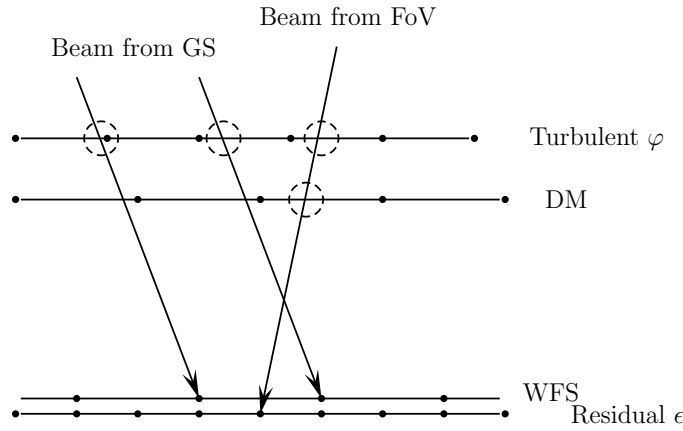


Figure 4. Typical geometry using influence functions.

The minimum variance reconstructor minimizes $\langle \|\epsilon\|^2 \rangle$ where $\langle \cdot \rangle$ indicates averaging over turbulence and noise statistics, and $\|\epsilon\|^2$ is the normalized *aperture-plane, piston removed* mean square wavefront error (mse) (overall

measure of performance) remaining after the estimated commands \hat{a} have been applied to the DM's, averaged over \bar{N} observation directions within the FoV ($\bar{N} \geq N$) with weights $\{\omega_{\bar{n}}\}_{1 \leq \bar{n} \leq \bar{N}}$:

$$\|\epsilon\|^2 \stackrel{\text{def}}{=} \epsilon^T \mathcal{W} \epsilon / (H^\varphi \varphi)^T \mathcal{W} (H^\varphi \varphi), \quad (2.8)$$

where ϵ is a concatenation of \bar{N} subvectors $\epsilon_{\bar{n}}$ corresponding to the different observation directions. The weighting matrix $\mathcal{W} = \text{Diag}(\omega_{\bar{n}} W)_{1 \leq \bar{n} \leq \bar{N}}$ accounts for aperture edge effects to obtain a more accurate residual wavefront variance, and must define a *positive definite metric* found by setting the norm squared of any aperture-plane vector u , $\|u\|_W^2 = u^T W u$, equal to norm squared of the corresponding continuous function as given by (2.5):

$$\|u\|_W^2 = \int dx dy \mathcal{A}(x, y) \left[u^I(x, y) - \int dx' dy' \mathcal{A}(x', y') u^I(x', y') \right]^2, \quad (2.9)$$

$\mathcal{A}(x, y)$ being the normalized aperture function ($\int \mathcal{A}(x, y) dx dy = 1$). The entries of W are then easily identified as

$$\begin{aligned} W &= W_0 - v_1 v_1^T, \quad (\text{sparse} + \text{low rank}) \\ (W_0)_{ij} &= \int dx dy \mathcal{A}(x, y) e_i(x, y) e_j(x, y), \\ (v_1)_i &= \int dx dy \mathcal{A}(x, y) e_i(x, y). \end{aligned} \quad (2.10)$$

The overall weighting matrix \mathcal{W} may thus be decomposed as a sparse plus low-rank matrix:

$$\begin{aligned} \mathcal{W} &= \mathcal{W}_0 - \mathcal{W}_1 \mathcal{W}_1^T, \\ \mathcal{W}_0 &= \text{Diag}(\omega_{\bar{n}} W_0)_{1 \leq \bar{n} \leq \bar{N}}, \\ \mathcal{W}_1 &= \text{Diag}(\omega_{\bar{n}}^{1/2} v_1)_{1 \leq \bar{n} \leq \bar{N}}. \end{aligned} \quad (2.11)$$

The pupil-plane residual piston-removed wavefront $\epsilon_{\bar{n}}$ in direction \bar{n} is given by

$$\begin{aligned} \epsilon_{\bar{n}} &= (H_{\bar{n}l}^\varphi)_{1 \leq l \leq L} \varphi - (H_{\bar{n}m}^a)_{1 \leq m \leq M} \hat{a} \\ &= \sum_{l=1}^L \varphi_{\bar{n}l}^I - \sum_{m=1}^M \hat{a}_{\bar{n}m}^I, \end{aligned} \quad (2.12)$$

where $H_{\bar{n}m}^a$ is the interpolation matrix for DM m in direction $\theta_{\bar{n}}$, whose columns are the influence functions $e_i(x)$ associated with each grid point i of DM m at altitude h_m , evaluated at $x = (\gamma_m x + \theta_{\bar{n}} h_m)$. The influence function $e_i(x)$ models the response of the i th actuator of DM m to a unit command.

The minimum variance reconstructor can be decomposed as

$$\mathcal{R} = F \mathcal{E}. \quad (2.13)$$

The estimated DM commands \hat{a} can then be computed in two steps as follows:

$$\hat{\varphi} = \mathcal{E} \begin{pmatrix} s \\ \tilde{s} \end{pmatrix}, \quad \hat{a} = F \hat{\varphi}, \quad \mathcal{E} = (\mathcal{E}_{ln})_{\substack{1 \leq l \leq L \\ 1 \leq n \leq N}}, \quad F = (F_{ml})_{\substack{1 \leq m \leq M \\ 1 \leq l \leq L}}. \quad (2.14)$$

The estimation matrix \mathcal{E} provides an estimate $\hat{\varphi}$ of the true turbulence profile φ , and the fitting matrix F provides a best fit of the actuator commands \hat{a} from the estimate $\hat{\varphi}$.

3. MINIMUM VARIANCE ESTIMATION MATRIX

The minimum variance estimation (MVE) matrix \mathcal{E} is formally given by

$$\begin{aligned} \mathcal{E} &= B^{-1} \mathcal{G}^T C^{-1}, \quad C^{-1} = \text{Diag}(C^{-1}, \tilde{C}^{-1}), \\ B &= \mathcal{G}^T C^{-1} \mathcal{G} + \mathcal{L}^{-1}, \quad \mathcal{L}^{-1} = \text{Diag}(\mathcal{L}_l^{-1})_{1 \leq l \leq L}, \end{aligned} \quad (3.1)$$

where C , \tilde{C} are the diagonal sensor noise covariance matrices corresponding respectively to the high-order and full-aperture tip/tilt WFS's, and \mathcal{L} is the block diagonal atmospheric covariance matrix (different turbulent layers are statistically independent). The first step of the estimation may be computed as

$$\begin{pmatrix} u \\ \tilde{u} \end{pmatrix} = \mathcal{G}^T C^{-1} \begin{pmatrix} s \\ \tilde{s} \end{pmatrix} = \begin{pmatrix} G^T C^{-1} (I - P) s \\ \tilde{G}^T \tilde{C}^{-1} \tilde{s} \end{pmatrix}, \quad (3.2)$$

where \tilde{G} and \tilde{C} are low-rank matrices. The second and final step in the evaluation, is to efficiently solve the system

$$B \hat{\varphi} = \begin{pmatrix} u \\ \tilde{u} \end{pmatrix}. \quad (3.3)$$

To proceed with this computation, we decompose the matrix B into layered blocks:

$$\begin{aligned} B &= \mathcal{M} - U U^T + V V^T, \\ \mathcal{M} &= G^T C^{-1} G + \mathcal{L}^{-1}. \end{aligned} \quad (3.4)$$

where U and V are low-rank matrices ($\text{rank}(U) = 2H$, $\text{rank}(V) = 2T$) with layered block representation

$$\begin{aligned} U &= (U_l)_{1 \leq l \leq L}, \quad U_l = ((G_{hl})^T \mathcal{Q}_h)_{1 \leq h \leq H}, \quad \mathcal{Q}_h = C_h^{-1} K_h (K_h^T C_h^{-1} K_h)^{-1/2}, \\ V &= (V_l)_{1 \leq l \leq L}, \quad V_l = \left((\tilde{G}_{tl})^T \tilde{C}_t^{-1/2} \right)_{1 \leq t \leq T}. \end{aligned} \quad (3.5)$$

The layered decomposition of B then reads

$$B_{ll'} = \mathcal{M}_{ll'} - \sum_{h=1}^H (H_{hl}^\varphi)^T \Gamma_h^T C_h^{-1} P_h \Gamma_h H_{hl'}^\varphi + \sum_{t=1}^T (\tilde{H}_{tl}^\varphi)^T \tilde{\Gamma}_t^T \tilde{C}_t^{-1} \tilde{\Gamma}_t \tilde{H}_{tl'}^\varphi. \quad (3.6)$$

3.1. Turbulence Covariance Matrix Approximation

Assuming that turbulence is stationary (i.e. the autocorrelation function depends only on the position difference between the pairs of positions involved), the atmospheric phase covariance matrix for layer l , \mathcal{L}_l , is a block Toeplitz with Toeplitz blocks (BTTB) matrix. Unfortunately, there is no simple formula for the inverse of BTTB matrices. Therefore, such matrices are often approximated by block circulant with circulant blocks (BCCB) matrices. This approximation is equivalent to a periodic extension of the phase screen. Consequently, on a finite grid, the BCCB approximation introduces a low-frequency error which will be negligible provided the computational domain is large compared to the telescope diameter.

The BCCB approximation allows us to easily invert the covariance matrix. Indeed, BCCB matrices are diagonalized by the 2-D discrete Fourier transform (DFT), so we can write

$$\mathcal{L}_l \simeq \mathcal{F}^{-1} \Lambda_l \mathcal{F}, \quad (3.7)$$

where \mathcal{F} is the 2-D DFT matrix. A good approximation to \mathcal{L}_l^{-1} , which appears in the reconstructor matrix (3.1), is then

$$\mathcal{L}_l^{-1} \simeq \mathcal{Z}_l \stackrel{\text{def}}{=} \mathcal{F}^{-1} \Lambda_l^{-1} \mathcal{F}. \quad (3.8)$$

This matrix is also BCCB. The diagonal entries of Λ come from the von Karman spatial power spectral density (PSD) of the atmospheric refractive index fluctuations δn , with the universal $-11/3$ power law:

$$\Lambda_l(\kappa) = \langle |\delta n(\kappa)|^2 \rangle = c_l^2 / [|\kappa|^2 + 1/L_0^2]^{11/6}. \quad (3.9)$$

Since $\Lambda_l(\kappa)$ is positive, \mathcal{Z}_l is also symmetric positive definite (SPD). Here κ is the spatial frequency vector (in rads/meter), L_0 is the turbulence outer-scale, and c_l is a constant defining the strength of layer l . For the case of a Kolmogorov turbulence spectrum, $L_0 = \infty$.

\mathcal{Z}_l is a non-sparse full rank matrix. As suggested in Ref.1, this matrix can be approximated by \mathcal{S}_l^2 , a scaled discrete Laplacian squared matrix (also known as biharmonic operator) with periodic boundary. Indeed, the discrete Laplacian \mathcal{S} has the Fourier representation

$$\hat{v} = 4 \left[\sin^2(\pi\kappa_x \Delta x) / \Delta x^2 + \sin^2(\pi\kappa_y \Delta y) / \Delta y^2 \right] \xrightarrow{\Delta x, \Delta y \rightarrow 0} 4\pi^2(\kappa_x^2 + \kappa_y^2) = 4\pi^2|\kappa|^2. \quad (3.10)$$

The advantage of this approximation is that the discrete Laplacian $\mathcal{S}_l = c_l' \mathcal{S}$ is sparse, enabling a sparse analog of the estimation matrix. If we adopt this approximation, the blocks of \mathcal{M} in the estimation matrix are given by

$$\mathcal{M}_{ll'} = \sum_{h=1}^H (G_{hl})^T C_h^{-1} G_{hl'} + \mathcal{S}_l^2 \delta_{ll'}, \quad 1 \leq l, l' \leq L. \quad (3.11)$$

We expect this approximation to become less satisfactory as the AO system increases.

4. MINIMUM VARIANCE FITTING MATRIX

The minimum variance fitting (MVF) matrix is formally given by

$$\begin{aligned} F &= A^{-1} H^{aT} \mathcal{W} H^a, \\ A &= H^{aT} \mathcal{W} H^a + V V^T + \alpha I, \end{aligned} \quad (4.1)$$

where V is a low-rank matrix whose columns are independent vectors from the null space of $H^{aT} \mathcal{W} H^a$, including for each DM layer, piston and cancelling tip/tilt modes, which have no effect on the piston-removed residual wavefront variance. α is a positive regularization parameter that must be included for MCAO systems to account for the modes in the null space that are not known a priori. The first step of the fitting part of the reconstructor can be computed easily from the sparse plus low-rank decomposition of the weighting matrix \mathcal{W} . To solve the fitting system, we proceed along the same lines as for the estimation part, i.e we decompose A into layered blocks:

$$\begin{aligned} A &= \mathcal{M} - U U^T + V V^T, \\ \mathcal{M} &= H^{aT} \mathcal{W}_0 H^a + \alpha I, \\ U &= H^{aT} \mathcal{W}_1 = (U_{m\bar{n}})_{\substack{1 \leq m \leq M \\ 1 \leq \bar{n} \leq \bar{N}}} , \quad U_{m\bar{n}} = (H_{\bar{n}m}^a)^T \omega_{\bar{n}}^{1/2} v_1, \\ V &= (V_m)_{1 \leq m \leq M}, \end{aligned} \quad (4.2)$$

where $\text{rank}(V) = 2$ (piston, tip-tilt), $\text{rank}(U) = \bar{N}$. The layered decomposition of A then reads

$$\begin{aligned} A_{ll'} &= \mathcal{M}_{ll'} - \sum_{\bar{n}=1}^{\bar{N}} U_{l\bar{n}} (U_{l'\bar{n}})^T + V_l V_{l'}^T = \mathcal{M}_{ll'} - \sum_{\bar{n}=1}^{\bar{N}} [(H_{\bar{n}l}^a)^T \omega_{\bar{n}} v_1] [v_1^T H_{\bar{n}l'}^a] + V_l V_{l'}^T, \\ \mathcal{M}_{ll'} &= \sum_{\bar{n}=1}^{\bar{N}} (H_{\bar{n}l}^a)^T \omega_{\bar{n}} W_0 H_{\bar{n}l'}^a + \alpha I_l \delta_{ll'}. \end{aligned} \quad (4.3)$$

5. MG LAYER-ORIENTED PCG

From the above discussion, both MVE and MVF matrix systems require the solution of symmetric positive definite (SPD) block matrix systems of generic form

$$Ax = b, \quad (5.1)$$

where x is a concatenated vector whose support is either the set of atmospheric or DM layers. In both cases, the matrix A can be represented in the form :

$$\begin{aligned} A &= \mathcal{M} - U U^T + V V^T = \mathcal{M} - \mathcal{U} \mathcal{V}^T, \\ \mathcal{U} &= (U \ V), \quad \mathcal{V} = (U \ -V). \end{aligned} \quad (5.2)$$

where U and V are low-rank full matrices. In order to solve (5.1), we call the Sherman-Morrison-Woodbury formula to invert low-rank perturbations to a matrix. The formula (also known as matrix inversion lemma) reads:

$$(\mathcal{M} - \mathcal{U}\mathcal{V}^T)^{-1} = \mathcal{M}^{-1} + \mathcal{M}^{-1}\mathcal{U} (I - \mathcal{V}^T \mathcal{M}^{-1}\mathcal{U})^{-1} (\mathcal{M}^{-1}\mathcal{V})^T. \quad (5.3)$$

Note that $\text{rank}(I - \mathcal{V}^T \mathcal{M}^{-1}\mathcal{U}) = \text{rank}(U) + \text{rank}(V)$. For natural guide star (NGS) MCAO systems, U and V are null matrices for the estimation step. A simplified (non-block) version for the single-layer case has already been solved efficiently at a cost of $\mathcal{O}(n \log n)$ using a multigrid preconditioned conjugate gradient method.² We propose to develop block extensions for the case of multiple layers.

The conjugate gradient (CG) method⁶ is an iterative scheme for solving SPD linear systems (5.1). Starting with $x_0 = 0$, CG generates polynomial approximates $x_\nu = -P_\nu(A)b$. The polynomials P_ν are optimal in the sense that they minimize $\|e_\nu\|_A^2$ where $e_\nu = x_\nu - x^*$ and $\|e_\nu\|_A^2 = e_\nu^T A e_\nu$. If A has q distinct eigenvalues, CG is guaranteed to converge to the exact solution x^* in at most q iterations, and it is “matrix free” in the sense that it does not require explicit storage of the matrix A . Each iteration of CG requires one matrix-vector product $z = Av$. In AO applications, A is very poorly conditioned, and this leads to very slow convergence of standard methods like CG when applied directly to (5.1). Therefore, the required number of iterations may be unacceptably high. Preconditioning⁶ refers to a technique to speed up the convergence of iterative methods. A preconditioner A' is an SPD approximation of A such that $\text{cond}(A'^{-1}A) \ll \text{cond}(A)$, i.e. $A'^{-1}A$ has a more desirable eigenvalue structure (clustering of the eigenvalues and/or low condition number) leading to rapid convergence of CG. The preconditioned conjugate gradient (PCG) algorithm solves the equivalent system

$$A'^{-1}Ax = A'^{-1}b. \quad (5.4)$$

Application of the preconditioner requires the computation of vector matrix multiplies of the form $u = A'^{-1}v$. We propose a multigrid algorithm to implement this preconditioning step.

The PCG algorithm has the following standard form:

```

ν := 0; φ₀      % initial guess
r₀ := Ax₀ - b; % initial residual
y₀ := A'⁻¹ r₀; % apply preconditioner
p₀ := -y₀;     % initial search direction
δ₀ := r₀ᵀ y₀;
Begin PCG Iterations
  h_ν := Ap_ν;
  τ_ν := δ_ν / p_νᵀ h_ν; % line search parameter
  x_{ν+1} := x_ν + τ_ν p_ν; % update approximate solution
  r_{ν+1} := r_ν + τ_ν h_ν; % update residual
  y_{ν+1} := A'⁻¹ r_{ν+1}; % apply preconditioner
  δ_{ν+1} := r_{ν+1}ᵀ y_{ν+1};
  β_{ν+1} := δ_{ν+1} / δ_ν;
  p_{ν+1} := -y_{ν+1} + β_{ν+1} p_ν; % update search direction
  ν := ν + 1;
end PCG iterations

```

5.1. The Multigrid Preconditioner

Multigrid (MG) methods⁷ are extremely effective for solving strongly elliptic partial differential equations (PDE's) (e.g. Laplace or biharmonic equations). The MVE matrix (3.11) falls into this class, and MG has been demonstrated to be very effective to solve (5.1) for ExAO systems.² MG is a multiresolution method. The key idea of a 2-grid scheme is to apply an iterative method (smoother) with fast convergence for high frequencies, project the (low-frequency) error onto a coarser grid (restriction), solve the system on the coarse grid, project back the coarse grid solution error onto the fine grid (interpolation), update the solution on the

fine grid and reapply the smoother to get rid of the new high-frequencies introduced. Multigrid methods are a recursive application of the 2-grid idea. To solve $A'u = v$, MG requires the following components:

- (i) A nested sequence of grids where $e_h = A_h^{-1}r_h$ is defined.
- (ii) Restriction and interpolation operators (transpose of each other to preserve symmetry for CG).
- (iii) A smoother: a stationary iterative method (characterized by an iteration matrix) which rapidly damps out high frequencies of the residual at each grid level.

A typical MG V-cycle is provided below.

```

Procedure  $x_h := \text{Vcycle}(level, x_h, b_h)$ 
Get  $A_h$ ;
If  $level = 1$  then
     $x_h := A_h^{-1}b_h$ ;    % Exact solution on coarsest grid
Else
     $x_h := S_h(B_h, x_h, b_h)$ ;    % Apply smoother
     $r_h := b_h - A_h x_h$ ;    % Compute residual
     $r_H := \text{Restrict}(r_h)$ ;    % Restrict residual to coarse grid
     $level = level - 1$ ;
     $e_H := 0$ ;
     $e_H := \text{Vcycle}(level, e_H, r_H)$ ;    %  $e_H \approx A_H^{-1}r_H$ 
     $e_h = \text{Interp}(e_H)$ ;    % Project back to fine grid
     $level = level + 1$ ;
     $x_h := x_h + e_h$ ;    % Update fine grid solution
     $x_h := S_h(A_h, x_h, b_h)$ ;    % Apply smoother
Endif

```

5.2. The layer-oriented Symmetric Gauss-Seidel Smoother

The recent ‘‘layer-oriented’’ control algorithms for MCAO⁸⁻¹¹ suggest to use a layer-oriented preconditioner. The layer-oriented approach uses WFS’s conjugated to the DM’s instead of to the entrance pupil, and an optical scheme to combine the signals from all guide stars, which has the effect of diminishing the cross-coupling of the reconstructor, such that each WFS can control its corresponding DM directly. It is argued that such systems may be less complex and less demanding in terms of computation.

In our computational approach, the first step in the estimation matrix \mathcal{E} (i.e. $\mathcal{G}^T \mathcal{C}^{-1}$) numerically combines the measurements from multiple WFS’s in the same way as Ragazzoni’s optical scheme with pyramid wavefront sensors. Therefore, we propose a layer-oriented (block) symmetric Gauss-Seidel (SGS) iteration as a smoother for MG. The iteration is based on the following partitioning of A :

$$A = A_D - A_L - A_U, \quad (5.5)$$

where A_D, A_L and A_U are block matrices assembled from the diagonal, lower and upper blocks of A respectively. The SGS iteration reads

$$\begin{aligned} x^{(2k-1)} &= (A_D + A_L)^{-1} y^{(2k-2)}, & x^{(2k)} &= (A_D + A_U)^{-1} y^{(2k-1)}, \\ y^{(2k-2)} &\stackrel{\text{def}}{=} b - A_U x^{(2k-2)}, & y^{(2k-1)} &\stackrel{\text{def}}{=} b - A_L x^{(2k-1)}. \end{aligned} \quad (5.6)$$

For layer j ($1 \leq j \leq j_{\max}$) where $j_{\max} = L, M$ for the estimation and fitting steps respectively, these read

$$\begin{aligned} y_{j_{\max}}^{(2k-2)} &= b_{j_{\max}}^{(2k-2)}, & y_j^{(2k-2)} &= b_j - \sum_{j'=j+1}^{j_{\max}} A_j j' x_{j'}^{(2k-2)}, \\ y_1^{(2k-1)} &= b_1^{(2k-1)}, & y_j^{(2k-1)} &= b_j - \sum_{j'=1}^{j-1} A_j j' x_{j'}^{(2k-1)}. \end{aligned} \quad (5.7)$$

Note that the iteration preserves cross-coupling between layers and can be implemented using back-substitutions:

$$\begin{aligned} x_j^{(2k-1)} &= A_{jj}^{-1} \left[y_j^{(2k-2)} - \sum_{j'=1}^{j-1} A_{jj'} x_{j'}^{(2k-1)} \right], \\ x_j^{(2k)} &= A_{jj}^{-1} \left[y_j^{(2k-1)} - \sum_{j'=j+1}^{j_{\max}} A_{jj'} x_{j'}^{(2k)} \right], \end{aligned} \quad (5.8)$$

which requires inversion of only diagonal blocks

$$A_{jj} = \mathcal{M}_{jj} - U_j U_j^T + V_j V_j^T, \quad 1 \leq j \leq j_{\max}. \quad (5.9)$$

This is done again using the matrix inversion lemma (5.3) implemented with Cholesky factorizations of the diagonal blocks \mathcal{M}_{jj} (complete or incomplete) with much less fill-in than factorization of the full matrix \mathcal{M} since off-diagonal blocks have been eliminated.

6. NUMERICAL SIMULATION RESULTS

We have arrived at a general framework for Minimum Variance Reconstructors, with a variety of options for special cases. The basic approach for the estimation step is the multigrid-preconditioned CG with a layer-oriented (block) SGS smoothing. Using a sparse representation for the turbulence inverse covariance matrix, diagonal blocks of the SGS smoother can be inverted using a symmetric approximate minimum degree (SYMAMD) permutation prior to Cholesky factorization to avoid fill-in. For the fitting step, the basic approach is incomplete Cholesky-preconditioned CG. Low rank perturbations for LGS and tip-tilt NGS measurements are easily accounted for using the matrix inversion lemma. Exact representation of the turbulence statistics requires an inner iterative solver like multigrid to invert the diagonal blocks \mathcal{M}_{jj} , which greatly increases the overall computational cost.

Figure 5 shows typical convergence results for the estimation step for telescope diameters equal to 32m, 16m and 8m. We consider the following 6-layer atmospheric profile obtained using generalized SCIDAR measurements at Cerro Pachon, Chile:

Layer	Altitude (m)	Relative weight
1	0	0.652
2	2580	0.172
3	5160	0.055
4	7730	0.025
5	12890	0.074
6	15460	0.022

The MCAO system under consideration has $N = 5$ NGS's at the edges and center of a FoV of diameter 2 arcmin. We are considering an array of 5×5 observation directions partitioning the FoV. The weights $\omega_{\bar{n}}$ of the weighting matrix \mathcal{W} in (2.8) have been chosen each equal to $1/\bar{N} = 1/25$. Our WFS sampling is at half phase screen resolution. The estimation error norm averaged over the FoV has been computed as

$$\begin{aligned} \|\epsilon_{\text{est}}\|^2 &\stackrel{\text{def}}{=} \epsilon_{\text{est}}^T \mathcal{W} \epsilon_{\text{est}} / (H^\varphi \varphi)^T \mathcal{W} (H^\varphi \varphi), \\ \epsilon_{\text{est}} &= H^\varphi (\hat{\varphi} - \varphi). \end{aligned} \quad (6.1)$$

Comparison with direct matrix inversion has been made using the following error metric:

$$\begin{aligned} \|\epsilon_{\text{ex}}\|^2 &\stackrel{\text{def}}{=} \epsilon_{\text{ex}}^T \mathcal{W} \epsilon_{\text{ex}} / (H^\varphi \hat{\varphi}_{\text{ex}})^T \mathcal{W} (H^\varphi \hat{\varphi}_{\text{ex}}), \\ \epsilon_{\text{ex}} &= H^\varphi (\hat{\varphi} - \hat{\varphi}_{\text{ex}}). \end{aligned} \quad (6.2)$$

Finally, the rms estimation error averaged over the FoV displayed in the lower-right panel of Figure 5 has been computed as

$$\xi_{\text{est}} = \left(\frac{\sum_{\bar{n}=1}^{\bar{N}} \omega_{\bar{n}} (H^\varphi \hat{\varphi} - H^\varphi \varphi)_{\bar{n}}^2}{\sum_{\bar{n}=1}^{\bar{N}} \omega_{\bar{n}} \| (H^\varphi \varphi)_{\bar{n}} \|^2} \right)^{1/2}, \quad (6.3)$$

where component-wise operation has to be understood for the square appearing in the numerator and the overall square root. We used 1 Vcycle/CG iteration and 1 SGS iteration/grid level. r_0 is equal to 25cm and the phase screens resolution is $dx = r_0$.

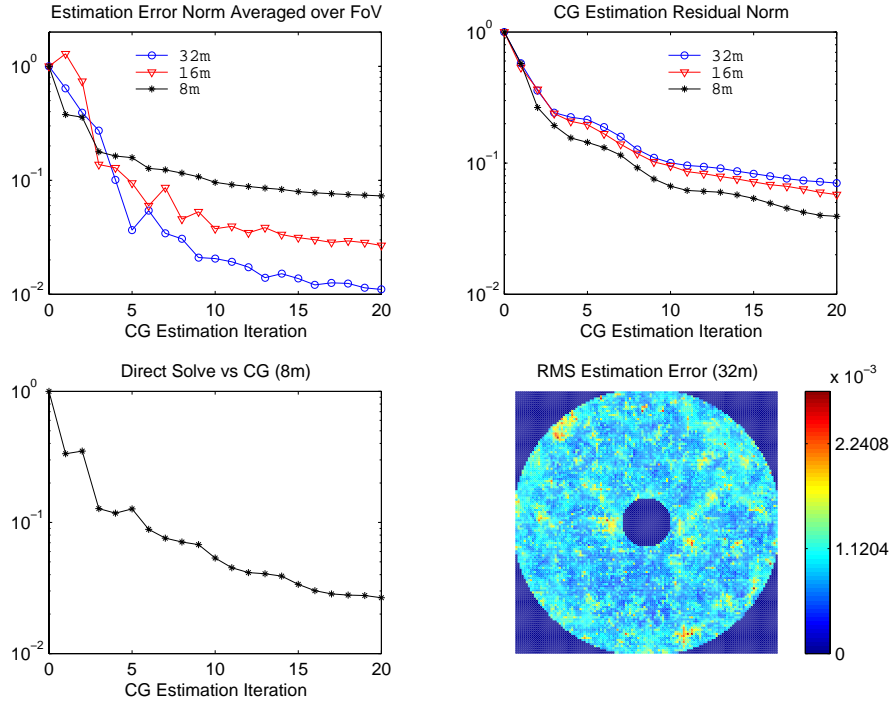


Figure 5. MVE for the Cerro-Pachon 6-layer profile, telescopes diameters equal to 32m, 16m and 8m. 5 WFSs' using 5 NGS's at the edges and center of a FoV of diameter 2 arcmin, 1 Vcycle/CG iteration, 1 SGS iteration/grid level, SNR = 20, $r_0 = 25$ cm, $dx = r_0$.

Figure 6 compares convergence results for the fitting step as the number of DM's varies from 1 to 3. We are considering the following 3 different DM configurations:

DM	Altitude (m)
1	0
2	5160
3	10320

3-DM Configuration

DM	Altitude (m)
1	0
2	2580

2-DM Configuration

DM	Altitude (m)
1	0

1-DM Configuration

All DM's are at half the phase screens resolution. The lower-left panel of Figure 6 displays the residual error after 2 CG fitting iterations for the 3-DM configuration as a function of CG estimation iterations. Figure 7 shows CG residual norm convergence for the 8m and 16m cases. Comparison with direct matrix inversion is presented in the upper-left panel for the 8m case. The rms residual error averaged over the FoV for the 32m, 2-DM case, displayed in the lower-right panel of Figure 7 has been computed as

$$\xi_{\text{fit}} = \left(\frac{\sum_{\bar{n}=1}^{\bar{N}} \omega_{\bar{n}} (H^a \hat{a} - H^\varphi \varphi)_{\bar{n}}^2}{\sum_{\bar{n}=1}^{\bar{N}} \omega_{\bar{n}} \| (H^\varphi \varphi)_{\bar{n}} \|^2} \right)^{1/2}, \quad (6.4)$$

where again component-wise operation has to be understood for the square appearing in the numerator and the overall square root.

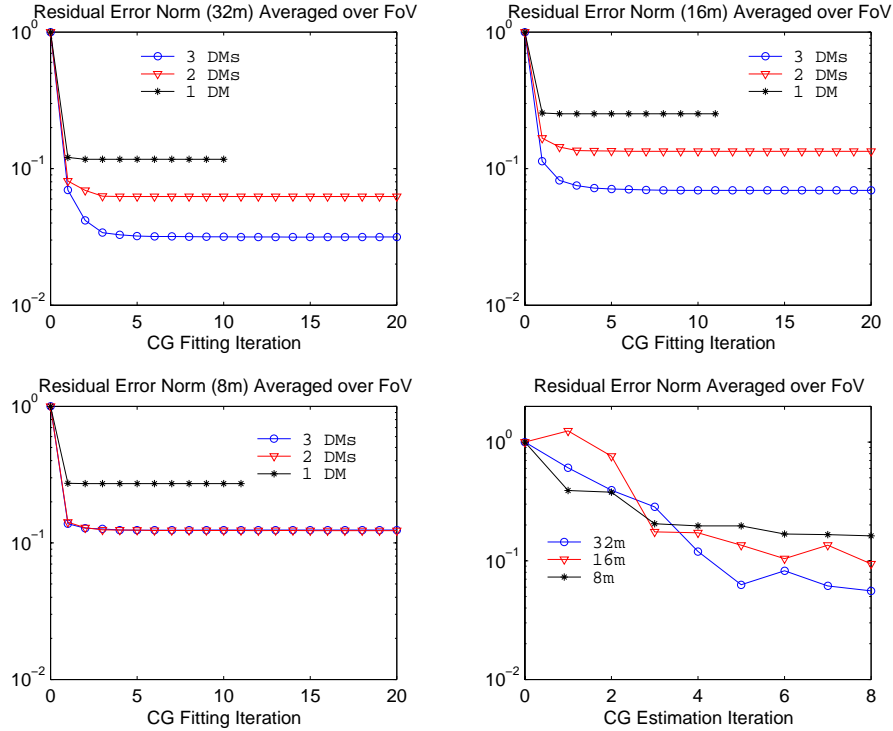


Figure 6. MVF for the same system as in Fig.5 with 3 different DM configurations. Averaging over an array of 5×5 observation directions partitioning a FoV of diameter 2 arcmin. Incomplete Cholesky preconditioning, regularization parameter $\alpha = 10^{-5}$, SNR = 20, $r_0 = 25$ cm, $dx = r_0$.

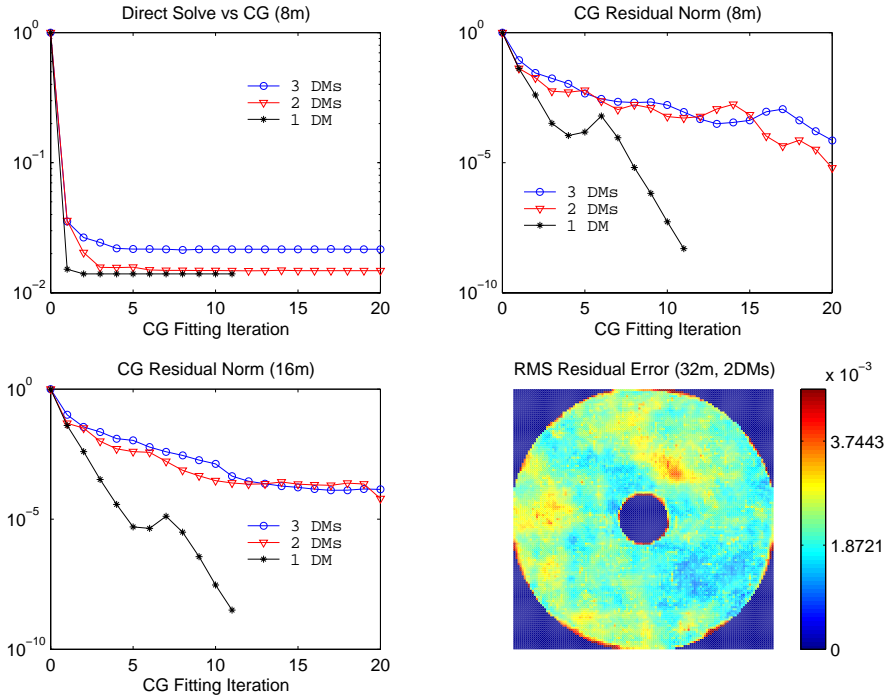


Figure 7. Residual Error after the DM's commands have been applied as a function of CG fitting iterations for the same system as in Figures 5 and 6.

7. CONCLUSIONS

We have developed efficient layer-oriented multigrid wavefront reconstruction algorithms for open-loop multi-conjugate adaptive optics systems on 32m, 16m and 8m telescopes. For the estimation step, the basic approach is the conjugate gradient method with a multigrid preconditioner to speed up convergence, and a layer-oriented (block) symmetric Gauss-Seidel smoother. Using a sparse representation for the turbulence inverse covariance matrix, diagonal blocks of the SGS smoother can be inverted using a symmetric approximate minimum degree (SYMAMD) permutation prior to Cholesky factorization to avoid fill-in. For the fitting step, the basic approach is incomplete Cholesky-preconditioned CG. Low rank perturbations for LGS and tip-tilt NGS measurements are easily accounted for using the matrix inversion lemma. Exact representation of the turbulence statistics requires an inner iterative solver like multigrid to invert the diagonal blocks, which greatly increases the overall computational cost.

ACKNOWLEDGMENTS

We acknowledge the support of the US Air Force Office of Scientific Research and the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini Partnership: the National Science Foundation (United States), the Particle Physics and Astronomy Research Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), CNPq (Brazil), and CONICET (Argentina).

REFERENCES

1. B. L. Ellerbroek, "Efficient computation of minimum variance wavefront reconstructors using sparse matrix techniques," *J. Opt. Soc. Am.* **A19**, September, 2002.
2. L. Gilles, C. R. Vogel and B. L. Ellerbroek, "A multigrid preconditioned conjugate gradient method for large scale wavefront reconstruction," *J. Opt. Soc. Am.* **A19**, September, 2002.

3. L. A. Poyneer, D. T. Gavel and J. M. Brase, “Fast wavefront reconstruction in large adaptive optics systems using the Fourier transform,” manuscript available from Lisa Poyneer, Lawrence Livermore National Laboratory, 2002.
4. B. L. Ellerbroek, “Scaling multi-conjugate adaptive optics performance estimates to extremely large telescopes,” *Proc. SPIE* **4007**, pp. 1088–1099, 2000.
5. O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computation*, SIAM, Philadelphia, 2001.
6. Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, Boston, 1996.
7. U. Trottenberg, C. W. Oosterlee and A. Schüller, *Multigrid*, Academic Press, London, 2001.
8. R. Ragazzoni, J. Farinato and E. Marchetti, “Adaptive optics for 100m class telescopes: new challenges require new solutions,” in *Beyond Conventional Adaptive Optics*, <http://lenin.pd.astro.it/venice2001/proceedings/>, *Proc. ESO*, 2001.
9. R. Ragazzoni, E. Diolaiti, J. Farinato, E. Fedrigo, E. Marchetti, M. Tordi and D. Kirkman, “Multiple Field of View Layer Oriented,” in *Beyond Conventional Adaptive Optics*, <http://lenin.pd.astro.it/venice2001/proceedings/>, *Proc. ESO*, 2001.
10. M. Tordi, R. Ragazzoni and E. Diolaiti, “Simulation of a Layer Oriented MCAO system,” in *Beyond Conventional Adaptive Optics*, <http://lenin.pd.astro.it/venice2001/proceedings/>, *Proc. ESO*, 2001.
11. A. A. Tokovinin, “Maximum separation between guide stars in atmospheric tomography,” in *Beyond Conventional Adaptive Optics*, <http://lenin.pd.astro.it/venice2001/proceedings/>, *Proc. ESO*, 2001.