Split atmospheric tomography using laser and natural guide stars

Luc Gilles* and Brent L. Ellerbroek

Thirty Meter Telescope Observatory Corporation, 1200 E. California Boulevard, Mail Code 102-8, Pasadena, California 91125
**Corresponding author: lgilles@caltech.edu*

Received May 22, 2008; revised July 15, 2008; accepted August 6, 2008; posted August 12, 2008 (Doc. ID 96540); published September 9, 2008

Laser guide star (LGS) atmospheric tomography is described in the literature as integrated minimum-variance tomographic wavefront reconstruction from a concatenated wavefront-sensor measurement vector consisting of many high-order, tip/tilt (TT)-removed LGS measurements, supplemented by a few low-order natural guide star (NGS) components essential to estimating the TT and tilt anisoplanatism (TA) modes undetectable by the TT-removed LGS wavefront sensors (WFSs). The practical integration of these NGS WFS measurements into the tomography problem is the main subject of this paper. A split control architecture implementing two separate control loops driven independently by closed-loop LGS and NGS measurements is proposed in this context. Its performance is evaluated in extensive wave optics Monte Carlo simulations for the Thirty Meter Telescope (TMT) LGS multiconjugate adaptive optics (MCAO) system, against the delivered performance of the integrated control architecture. Three iterative algorithms are analyzed for atmospheric tomography in both cases: a previously proposed Fourier domain preconditioned conjugate gradient (FDPCG) algorithm, a simple conjugate gradient (CG) algorithm without preconditioning, and a novel layer-oriented block Gauss–Seidel conjugate gradient algorithm (BGS-CG). Provided that enough iterations are performed, all three algorithms yield essentially identical closed-loop residual RMS wavefront errors for both control architectures, with the caveat that a somewhat smaller number of iterations are required by the CG and BGS-CG algorithms for the split approach. These results demonstrate that the split control approach benefits from (i) a simpler formulation of minimum-variance atmospheric tomography allowing for algorithms with reduced computational complexity and cost (processing requirements), (ii) a simpler, more flexible control of the NGS-controlled modes, and (iii) a reduced coupling between the LGS- and NGS-controlled modes. Computation and memory requirements for all three algorithms are also given for the split control approach for the TMT LGS AO system and appear feasible in relation to the performance specifications of current hardware technology. © 2008 Optical Society of America

OCIS codes: 010.1080, 010.1330.

1. INTRODUCTION

High-order, high-speed laser guide star (LGS) adaptive optics (AO) systems, with thousands of actuators updated at kilohertz rates, will play a key role in the development of the next generation of ground-based astronomical telescopes. Emerging LGS AO techniques for these telescopes include multiconjugate adaptive optics (MCAO), i.e., several LGSs and several deformable mirrors (DMs); laser tomography adaptive optics (LTAO), i.e., several LGSs and a single DM; and multiobject adaptive optics (MOAO), i.e., several LGSs and a single DM per field point. All of these AO modes require computationally efficient and parallelizable tomographic wavefront reconstruction algorithms to estimate the dominant layers of distributed volume turbulence above the telescope. First images of unprecedented quality were recently obtained with an MCAO instrument [1], demonstrating the maturity of the technology beyond simulations and laboratory experiments. Minimum-variance tomographic wavefront reconstruction followed by least-squares DM fitting is generally a requirement for these systems to achieve near optimal performance. However, explicit computation, storage, real-time update, and application of the reconstructor are prohibitively expensive. To overcome this problem, sparse-matrix techniques have been exploited [2], opening the door to iterative solutions for zonal LGS atmospheric tomography [3,4].

Current generation LGS AO systems [5] incorporate a low-order natural guide star (NGS) wavefront sensor (WFS) to measure the tip/tilt (TT) modes of the wavefront that are undetectable using a laser beacon owing its uncertain position [6]. Proposed MCAO systems will utilize a few low-order NGS WFSs to measure the TT and tilt anisoplanatism (TA) modes as well [7]. Since most TA occurs in the three quadratic Zernike modes [8], it can be controlled by two DMs and sensed by three low-order NGS WFSs [three TT WFSs or three tip/tilt/focus/astigmatism (TTFA), i.e., order 2×2 WFSs, or a mix of TT and TTFA WFSs]. In practical implementations, at least one TTFA WFS is required in order to measure defocus errors arising from variations in the range to sodium LGSs. The three TA modes can then be specified as focus and astigmatism distributed on two DMs, which are scaled so that the wavefront propagated from each one of the sources forming the LGS asterism (i.e., the array of LGS beacons) to the aperture plane consists of pure TT, which is invisible to the TT-removed LGS WFSs. These modes are sometime referred to as "plate scale modes" because

they introduce field-dependent TT errors that magnify and distort the image at the science instrument focal plane. They also introduce field-independent focus/astigmatism wavefront errors on account of the cone effect.

The LGS and NGS measurements have very different characteristics. The former are high order, are TT removed, have a relatively high and fixed signal-to-noise ratio (SNR), and are associated with a fixed guide star geometry. The latter are low order, associated with a variable guide star geometry, and generally have a low SNR. A separate NGS bandwidth optimization is therefore required, traditionally achieved by decomposition of the combined LGS/NGS wavefront reconstruction matrix into columns of actuator commands driven by LGS and NGS measurements. Such a decomposition becomes quickly impractical, however, as the order of the system increases and prevents explicit computation of the reconstructor.

The split atmospheric tomography wavefront-control architecture proposed in this paper for LGS MCAO is illustrated in the bottom panel of Fig. 1 and is a generalization of the architecture for current LGS AO systems [5]. A somewhat similar concept of split LGS/NGS MCAO wavefront reconstruction has already been developed for the Gemini South MCAO system [9]. The implementation proposed in this paper is based on two separate control loops driven independently by the LGS and NGS measurements. The NGS control loop uses a noise-weighted least-squares wavefront reconstruction matrix to control the two global TT modes and the three TA modes. The LGS control loop uses an iterative, computationally efficient algorithm providing an approximate solution to minimum-variance tomography, and it controls all modes orthogonal to the five NGS modes. The benefits of the split tomography approach are (i) a reduced coupling between the LGS- and NGS-controlled modes, (ii) a more flexible control of the TT and TA modes, and (iii) a simpler formulation of LGS tomography allowing for algorithms with reduced computational complexity and cost. These points will be further detailed below.

The paper is organized as follows. Sections 2 and 3 provide technical descriptions on the operations involved in, respectively, the integrated and split atmospheric tomography architectures. Sample simulation results for the Thirty Meter Telescope (TMT) LGS MCAO system are given in Section 4, and their computation and memory requirements are given in Section 5.

2. INTEGRATED ATMOSPHERIC TOMOGRAPHY

LGS atmospheric tomography consists of several computational steps that perform operations on three types of wavefront grids: a set of atmospheric grids, an aperture-plane grid, and DM grids. For computational convenience, all grids shall be square, with atmospheric grids scaled in a cone coordinate system where mesh size is squeezed with range according to the cone compression factor $\Delta_k = \xi_k \alpha_k \Delta_0$. Here $\xi_k = 1 - h_k/h_{\mathrm{lgs}}$ denotes the grid compression factor for layer k at range h_k and LGSs at range h_{lgs}, $\alpha_k = 1$ or 2 is the grid oversampling parameter, and $\Delta_0 = d_{\mathrm{sa}}/2$ denotes the aperture-plane grid mesh size, equal to half the LGS WFS subaperture size. A graphical illustration of the cone coordinate system is provided in Fig. 2. Note that regardless of the choice of coordinate system, any LGS tomographic wavefront reconstruction algorithm needs to be updated slowly with time owing to sodium layer range variations. In the cone coordinate system, this update is most conveniently accomplished by updating the atmospheric grids themselves, i.e., updating the above ξ_k parameters.

Denoting by N_{gs}, N_{lgs}, and N_{ngs} the total number of guide stars, LGSs, and low-order NGSs participating in the reconstruction of the atmospheric grids, we have $N_{\mathrm{gs}} = N_{\mathrm{lgs}} + N_{\mathrm{ngs}}$ for the integrated approach. Four core operations are involved in LGS atmospheric tomography:

- Bilinear interpolation. This operation computes wavefront values at the intercepts of rays traced through phase screens (atmospheric and DM grids) to/from the aperture plane, for sources at either finite or infinite range.

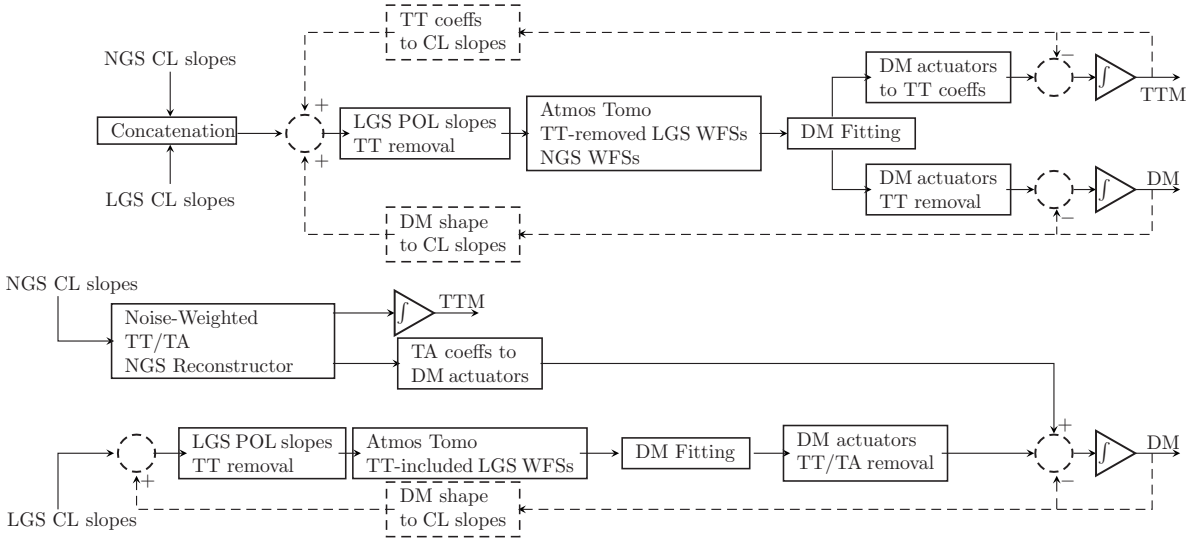

Fig. 1. Integrated (top) versus split (bottom) tomography block diagrams. Dashed lines indicate signals used to compute pseudo-open-loop (POL) measurements.

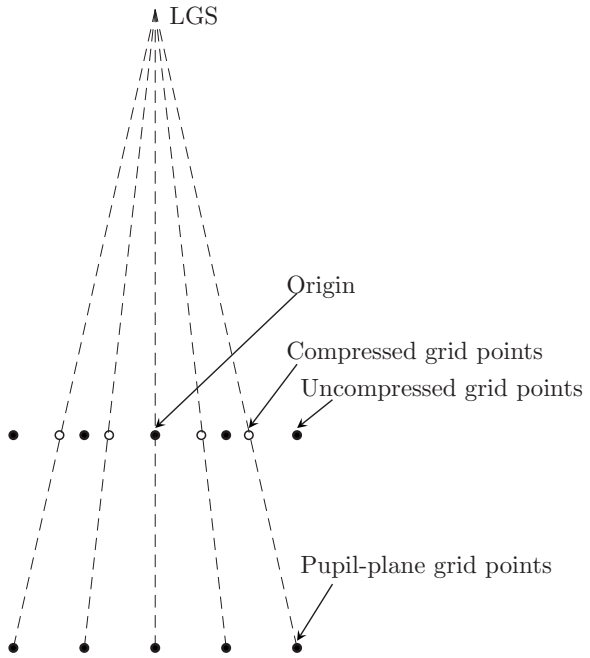

Fig. 2. Graphical representation of the ray-tracing operation for a finite-range point source, illustrating an upper altitude grid in either a compressed or an uncompressed coordinate system. The origin is common to both compressed and uncompressed grids. For an LGS off axis, the ray intercept offset is constant for the compressed grid.

Rays traced to/from the aperture plane correspond to forward/backward wavefront propagations, which are transposes of one another. The projected aperture-plane grid point coordinates on a plane at range h for a source at range h_{src} in direction θ are given by $x' = (1 - h/h_{\mathrm{src}})x + h\theta$, where x denotes aperture-plane grid point coordinates (see Fig. 2 for a graphical illustration). Note that atmospheric grids must be sufficiently large in order to properly sample all LGS and NGS beams.

- Aperture-plane LGS and NGS WFS gradient and gradient transpose. The gradient operator computes each WFS gradient measurement from aperture-plane wavefront values. These are sparse operations, coupling only aperture-plane grid points bordering opposite sides of a subaperture.
- Atmospheric wavefront curvature squared. This operation regularizes the tomography matrix system by providing a sparse approximation to the inverse atmospheric covariance matrix [2]. The curvature squared (biharmonic) operator couples 13 points lying on a 5×5 stencil, with 3 points from each of the 4 corners removed. The 13 weights are identical for all atmospheric phase screens and are identical throughout each screen, except near the boundary.
- Aperture weighting. This operation is performed in the DM fitting step to capture aperture edge effects and yield a lower fitting error by cross-coupling aperture-plane grid points with up to eight nearest neighbors on a regular 3×3 stencil. Fully interior stencils have two-dimensional Simpson weights.

The use of all four of these operations in the steps of atmospheric tomography are described in greater detail below. Grid-based as opposed to sparse-matrix-based operations are preferable for real-time implementation since they have the highest potential to fully exploit the underlying structure of these operators, minimize storage requirements, and, most importantly, exploit hardware parallelism. This is common practice in real-time image processing where bilinear interpolation and smoothing are frequently implemented [10].

A. Pseudo-Open-Loop Gradient Computation

Since atmospheric tomography uses an open-loop minimum-variance tomographic wavefront reconstructor in closed loop, it must operate on pseudo-open-loop (POL) gradients. This is also a requirement for closed-loop stability reasons. In the absence of measurement noise and systematic errors (WFS pupil distortion and misregistration/calibration errors), the poles of the closed-loop transfer matrix are then equal to those of a conventional minimum-variance reconstructor with closed-loop constraints [11]. POL gradients represent an estimate of the open-loop gradients that would have been measured with ideal flat DMs and ideal fully linear WFSs. As a result, the POL architecture eliminates the need for an invisible mode-removal step as in a conventional closed-loop system. Such a control approach has been implemented successfully in the past on laser beam control systems [12] and more recently in MCAO simulations [13]. An analysis of its temporal frequency response has also been reported [14].

The POL gradients associated with WFS number k are computed as follows:

$$[s_{\mathrm{dm}}]_k = \Gamma_k [[H^{\mathrm{gs}}_{\mathrm{dm}}]_{k1} \quad \cdots \quad [H^{\mathrm{gs}}_{\mathrm{dm}}]_{kN_{\mathrm{dm}}}] \begin{bmatrix} [\hat{x}]_1 \\ \vdots \\ [\hat{x}]_{N_{\mathrm{dm}}} \end{bmatrix}, \qquad 1 \leq k \leq N_{\mathrm{gs}}, \tag{1}$$

$$[s_{\mathrm{pol}}]_k = [s]_k + [s_{\mathrm{dm}}]_k, \tag{2}$$

where N_{dm} denotes the number of DMs; s_{pol} and s denote, respectively, the POL and closed-loop gradient vectors; $\hat{x}$ is the concatenated DM actuator vector; Γ_k denotes the aperture-plane gradient operator for WFS number k; and $[H^{\mathrm{gs}}_{\mathrm{dm}}]_{kl}$ represents the bilinear interpolation operator for the action of DM grid l upon rays traced from guide star number k to the aperture-plane grid. $[H^{\mathrm{gs}}_{\mathrm{dm}}]_{kl}$ will generally have the following stencil:

$$w = \begin{bmatrix} \beta_y \\ 1 - \beta_y \end{bmatrix} [1 - \beta_x \quad \beta_x] = \begin{bmatrix} (1 - \beta_x)\beta_y & \beta_x \beta_y \\ (1 - \beta_x)(1 - \beta_y) & \beta_x (1 - \beta_y) \end{bmatrix}, \tag{3}$$

where (β_x, β_y) is the ray intercept offset relative to the DM grid mesh size from the lower-left neighbor node. Note that this offset is not constant for LGSs, since DM grids are not in the cone coordinate system and the intercepts are from finite-range sources. The β_x/β_y offsets for each LGS can be stored separately in two one-dimensional arrays, since β_x is only a function of the x intercept of a ray in the aperture plane (and similarly for β_y). Finally, note that the values of β_x/β_y for the LGSs

need to be updated every few seconds as the range to the sodium layer varies.

The gradient operators Γ_k associated with fully illuminated subapertures have the following regular stencil:

$$w_x = \frac{1}{d_{\mathrm{sa}}}\begin{bmatrix} -1/2 & 0 & 1/2 \\ -1 & 0 & 1 \\ -1/2 & 0 & 1/2 \end{bmatrix}, \qquad w_y = -w_x^{\mathrm{T}}. \tag{4}$$

Note that the central point of the stencil has a zero weight for both measurement directions. For the gradient transpose operators (which will appear below), grid points that are the common vertex to four fully illuminated subapertures receive a nonzero contribution from both x and y gradients with the following weights:

$$w_x = \frac{1}{d_{\mathrm{sa}}}\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix}, \qquad w_y = -w_x^{\mathrm{T}}, \tag{5}$$

whereas grid points that are the common midpoint node bordering two fully illuminated subapertures receive a nonzero contribution from either the x or the y gradients with the following weights:

$$w_x = \frac{1}{d_{\mathrm{sa}}}[1 \quad -1], \qquad w_y = -w_x^{\mathrm{T}}. \tag{6}$$

B. Computing the Tomography System Right-Hand-Side Vector

Assembling the right-hand-side vector of the tomography matrix system from the POL gradients is accomplished by the following operations:

$$[y]_k = \Gamma_k^{\mathrm{T}} W_k \Theta_k [s_{\mathrm{pol}}]_k, \qquad 1 \leqslant k \leqslant N_{\mathrm{gs}} \tag{7}$$

$$[b]_l = [[H_{\mathrm{atm}}^{\mathrm{gs}}]_{1l}^{\mathrm{T}} \quad \cdots \quad [H_{\mathrm{atm}}^{\mathrm{gs}}]_{N_{\mathrm{gs}}l}^{\mathrm{T}}] \begin{bmatrix} [y]_1 \\ \vdots \\ [y]_{N_{\mathrm{gs}}} \end{bmatrix}, \qquad 1 \leqslant l \leqslant N_{\mathrm{ps}}, \tag{8}$$

where N_{ps} denotes the number of atmospheric grids, Θ_k denotes full-aperture noise-weighted gradient TT removal if WFS k corresponds to a LGS WFS and identity otherwise, W_k is the precomputed measurement noise inverse covariance matrix for WFS k, $[b]_l$ denotes the right-hand-side tomography subvector corresponding to phase screen l, and $[H_{\mathrm{atm}}^{\mathrm{gs}}]_{kl}$ is the bilinear interpolation operator between the nodes of atmospheric layer l and intercepts of rays traced to the aperture-plane grid from guide star k. Note that the interpolation offset (β_x, β_y) is constant throughout the grid for LGSs since grid nodes are defined in the cone coordinate system. This property is particularly well suited for a parallel grid-based implementation. The nodes of the aperture-plane grid can always be chosen to coincide with nodes of the ground-layer phase screen such that no interpolation is needed for that screen. W_k is a sparse symmetric weighting matrix with two nonzero elements per row (measurement noise between different subapertures is uncorrelated, but the x and y components of the noise in a subaperture may be correlated in the case of elongated LGSs).

C. Computing an Approximate Solution to the Tomography System

The tomography matrix system takes the form $Ax=b$, where x denotes the concatenated tomography vector of unknowns (atmospheric phase screens), A is the block structured tomography operator, and b is the tomography right-hand-side vector. All tomography solvers discussed below perform $[A]_{kl}[v]_l$ operations $(1 \leqslant k, l \leqslant N_{\mathrm{ps}})$, which can be expressed as follows:

$$\psi = \mathrm{Diag}(\Gamma_1^{\mathrm{T}} W_1 \Theta_1 \Gamma_1, \ldots, \Gamma_{N_{\mathrm{gs}}}^{\mathrm{T}} W_{N_{\mathrm{gs}}} \Theta_{N_{\mathrm{gs}}} \Gamma_{N_{\mathrm{gs}}}) \begin{bmatrix} [H_{\mathrm{atm}}^{\mathrm{gs}}]_{1l} \\ \vdots \\ [H_{\mathrm{atm}}^{\mathrm{gs}}]_{N_{\mathrm{gs}}l} \end{bmatrix} [v]_l, \tag{9}$$

$$[p]_k = [[H_{\mathrm{atm}}^{\mathrm{gs}}]_{1k}^{\mathrm{T}} \quad \cdots \quad [H_{\mathrm{atm}}^{\mathrm{gs}}]_{N_{\mathrm{gs}}k}^{\mathrm{T}}]\psi, \tag{10}$$

$$[q]_l = \gamma_l L^2 [v]_l, \tag{11}$$

$$[A]_{kl}[v]_l = [p]_k + [q]_l \delta_{kl}, \tag{12}$$

where δ_{kl} denotes the Kronecker delta, L^2 is the curvature squared (biharmonic) regularization operator coupling 13 grid points in a 5×5 stencil on each screen, and γ_l is a scaling constant proportional to Δ_l^{-4} and to the (5/6)th power of the Fried parameter of layer l.

The technique of warm restart has been shown in the past to significantly accelerate the closed-loop convergence of iterative wavefront reconstruction algorithms [15,16]. With this technique, the solution at a given frame is used as initial guess for the next frame instead of the null vector as in cold restart. Three warm-started iterative algorithms falling in two categories are proposed for solving the integrated tomography matrix system $Ax=b$: a layer-oriented algorithm operating on the blocks (i.e., layers) of the system one by one and two system-oriented algorithms operating on the whole system (i.e., all layers) at once. These algorithms are, respectively,

- BGS-CG: block Gauss–Seidel conjugate gradient. The block Gauss–Seidel (BGS) method is a block generalization of the well known Gauss–Seidel iteration that has been implemented in earlier MCAO simulations as smoothing step for the multigrid (MG) tomographic wavefront reconstructor [3]. It is classified in computational linear algebra as a domain decomposition method [17]. The coefficient matrix A is decomposed into a number of blocks $[A]_{kl}$ (associated with N_{ps} atmospheric layers in our case) and expressed in the form $A=A_{\mathrm{D}}+A_{\mathrm{L}}+A_{\mathrm{U}}$, where A_{D}, A_{L} and A_{U} are block structured matrices assembled from the diagonal, strictly lower, and strictly upper blocks of A respectively. Matrices A_{L} and A_{U} cross couple atmospheric layers, whereas the block diagonal matrix A_{D} couples phase grid points on individual layers. Starting with a warm initial guess $x^{(0)}$, a generic forward BGS iteration solves the following system:

$$(A_D + A_L)x = y, \tag{13}$$

$$y = b - A_U x^{(0)}. \tag{14}$$

Equation (13) defines a block lower triangular system that requires the solution of N_{ps} smaller and sparser systems $[A]_{kk}[x]_k = [v]_k$, where the right-hand-side vector $[v]_k$ depends on the solution obtained for all diagonal blocks above k. BGS-CG uses the conjugate gradient (CG) algorithm to obtain an approximate solution $[\hat{x}]_k$ of these N_{ps} smaller systems. Note that the CG solver alleviates the sequential processing (and hence the propagation delay) incurred by the Cholesky factorization implementation reported in [3].

The BGS-CG algorithm can thus be organized in two main steps: (i) assembling the transformed right-hand-side vector y and (ii) layer-by-layer solution of the block-structured system (13). These two steps can be formulated as follows:

$$[y]_{N_{ps}} = [b]_{N_{ps}}; \psi = 0$$

For $k = N_{ps} - 1, \ldots, 1$

$$\psi := \psi + \mathrm{Diag}(\Gamma_1^T W_1 \Theta_1 \Gamma_1, \ldots, \Gamma_{N_{gs}}^T W_{N_{gs}} \Theta_{N_{gs}} \Gamma_{N_{gs}}) \times \begin{bmatrix} [H_{atm}^{gs}]_{1k+1} \\ \vdots \\ [H_{atm}^{gs}]_{N_{gs}k+1} \end{bmatrix} [x^{(0)}]_{k+1}$$

$$[y]_k = [b]_k - [[H_{atm}^{gs}]_{1k}^T \quad \cdots \quad [H_{atm}^{gs}]_{N_{gs}k}^T]\psi$$

End

$$[\hat{x}]_1 = \mathrm{CG}([A]_{11}, [y]_1); \quad \psi = 0$$

For $k = 2, \ldots, N_{ps}$

$$\psi := \psi + \mathrm{Diag}(\Gamma_1^T W_1 \Theta_1 \Gamma_1, \ldots, \Gamma_{N_{gs}}^T W_{N_{gs}} \Theta_{N_{gs}} \Gamma_{N_{gs}}) \times \begin{bmatrix} [H_{atm}^{gs}]_{1k-1} \\ \vdots \\ [H_{atm}^{gs}]_{N_{gs}k-1} \end{bmatrix} [\hat{x}]_{k-1}$$

$$[v]_k = [y]_k - [[H_{atm}^{gs}]_{1k}^T \quad \cdots \quad [H_{atm}^{gs}]_{N_{gs}k}^T]\psi$$

$$[\hat{x}]_k = \mathrm{CG}([A]_{kk}, [v]_k)$$

End

where the function $\tilde{x} = \mathrm{CG}(A, b)$ performs warm-started CG iterations on the generic system $Ax = b$.

- CG: warm-started CG iterations without preconditioning operating on the whole tomography system.
- FDPCG: warm-started CG iterations with a Fourier domain preconditioning matrix [4]. The FDPCG algorithm operates in the spatial domain with a sparse precomputed Fourier domain Hermitian preconditioning matrix slowly updated as the range to the layers, range to the LGS, signal level, and turbulence conditions change. The preconditioning step can be implemented in parallel via N_{ps} fast Fourier transforms (FFTs), a Hermitian matrix vector multiplication, and N_{ps} inverse FFTs. Note that the originally proposed FDPCG algorithm [4] attempts to disentangle the LGS and NGS contributions by inappropriately invoking the matrix inversion lemma, which leads to closed-loop instabilities [15]. The above formulation avoids these problems by operating on the integrated tomography operator containing both LGS and NGS contributions.

Note that BGS-CG and CG have order N computational complexity, where N denotes the size of the tomography matrix system.

D. Propagating the Tomography Estimate

Once an approximate tomography solution has been computed, the tomography estimate is geometrically propagated (ray-trace interpolation/accumulation) along N_{fit} fitting directions from sources (science targets) at infinite range, which can be represented as follows:

$$[y]_k = [[H_{atm}^{sc}]_{k1} \quad \cdots \quad [H_{atm}^{sc}]_{kN_{ps}}] \begin{bmatrix} [\hat{x}]_1 \\ \vdots \\ [\hat{x}]_{N_{ps}} \end{bmatrix}, \qquad 1 \le k \le N_{fit}, \tag{15}$$

where $[\hat{x}]_l$ denotes the tomography estimate for layer l and $[H_{atm}^{sc}]_{kl}$ is a sparse bilinear interpolation operator between the nodes of atmospheric grid l and intercepts of rays traced to the aperture-plane grid from the fitting direction k. Note that the interpolation offsets (β_x, β_y) are not constant on a given phase screen, since atmospheric phase screens are in the cone coordinate system and ray tracing is from infinite-range sources. However, the offsets need to be stored for only one row/column of grid points for each infinite-range source.

E. Computing the Fitting System Right-Hand-Side Vector

The fitting right-hand-side vector is computed as follows:

$$[q]_k = \omega_k W [y]_k, \qquad 1 \le k \le N_{fit} \tag{16}$$

$$[b]_l = [[H_{dm}^{sc}]_{1l}^T \quad \cdots \quad [H_{dm}^{sc}]_{N_{fit}l}^T] \times \begin{bmatrix} [q]_1 \\ \vdots \\ [q]_{N_{fit}} \end{bmatrix}, \qquad 1 \le l \le N_{dm}, \tag{17}$$

where $[b]_l$ is the right-hand-side subvector for DM grid l, ω_k is the scalar weight for the fitting direction k, W is a fixed sparse aperture-plane weighting operator coupling each grid point with up to eight nearest neighbors on a 3×3 stencil (fully interior stencils have Simpson weights), and $[H_{dm}^{sc}]_{kl}$ is a sparse bilinear interpolation operator between the actuators of DM grid l and intercepts of the rays traced to the aperture-plane grid from the fitting direction k.

F. Computing an Approximate Solution to the Fitting System

The fitting matrix system to solve has again the form $Ax = b$, where x denotes the concatenated DM actuator vector of unknowns, b is the concatenated fitting right-

hand-side vector, and A is the block-structured fitting operator. CG (operating on the whole fitting system) without preconditioning is the proposed fitting solver. The solver performs $[A]_{kl}[v]_l$ operations ($1 \leq k,l \leq N_{\mathrm{dm}}$), which can be expressed as follows:

$$\psi = \mathrm{Diag}(\omega_1 W, \ldots, \omega_{N_{\mathrm{fit}}} W) \begin{bmatrix} [H_{\mathrm{dm}}]_{1l} \\ \vdots \\ [H_{\mathrm{dm}}]_{N_{\mathrm{fit}}l} \end{bmatrix} [v]_l, \quad (18)$$

$$[A]_{kl}[v]_l = [[H_{\mathrm{dm}}]_{1k}^{\mathrm{T}} \quad \cdots \quad [H_{\mathrm{dm}}]_{N_{\mathrm{fit}}k}^{\mathrm{T}}]\psi. \quad (19)$$

3. SPLIT ATMOSPHERIC TOMOGRAPHY

In split atmospheric tomography, only LGSs participate in the reconstruction of the atmospheric grids, i.e., $N_{\mathrm{gs}} = N_{\mathrm{lgs}}$. This leads to the following simplifications of the operations described in Section 2: (i) Atmospheric grid sizes no longer depend upon the variable location of the NGS asterism, (ii) ray tracing, aperture-plane gradient, and gradient transpose operations are no longer performed for the NGSs, and (iii) the full-aperture gradient TT-removal operator Θ_k is not applied inside the tomography operator, although it is applied to compute the right-hand-side tomography vector. On the other hand, following the DM fitting step, a NGS mode-removal step must be performed to project off the five TT/TA modes from the fitting estimate as described below.

Let us denote by M the five-column TT/TA modal matrix spanning the NGS actuator subspace, and by $G_{\mathrm{a}} = \Gamma^{\mathrm{ngs}} H_{\mathrm{dm}}^{\mathrm{ngs}}$ the DM-to-NGS WFS interaction matrix mapping DM actuators to geometric NGS WFS measurements. The proposed NGS reconstructor R mapping closed-loop NGS WFS measurements, s^{ngs}, to the error in the NGS DM actuator command, $\hat{e}$, is defined as the noise-weighted pseudoinverse of the modal DM-to-NGS WFS interaction matrix $G_{\mathrm{M}} = G_{\mathrm{a}} M$:

$$R = G_{\mathrm{M}}^{\dagger}, \quad (20)$$

$$G_{\mathrm{M}}^{\dagger} = (G_{\mathrm{M}}^{\mathrm{T}} W^{\mathrm{ngs}} G_{\mathrm{M}})^{-1} G_{\mathrm{M}}^{\mathrm{T}} W^{\mathrm{ngs}}, \quad (21)$$

where W^{ngs} is the inverse of the diagonal NGS measurement-noise covariance matrix. Such a reconstructor provides a noise-weighted least-squares fit of the closed-loop NGS WFS measurements to the gradient of the propagated error in the TT/TA modes along the NGS directions. The TT/TA mode coefficient error is then given by $\hat{e} = \arg\min_e \|s^{\mathrm{ngs}} - G_{\mathrm{M}} e\|_{W^{\mathrm{ngs}}}^2 = G_{\mathrm{M}}^{\dagger} s^{\mathrm{ngs}}$. The errors are then temporally filtered to produce the commands that are applied to the DMs and the TT mirror. Note that "ideal" mode coefficients would be given by the amplitude-weighted least-squares fit of the science wavefronts to the propagated TT/TA modes along the science directions. Such mode coefficients would be expressed as $\hat{m}_0 = \arg\min_m \|\phi^{\mathrm{sc}} - H_{\mathrm{M}}^{\mathrm{sc}} m\|_A^2 = H_{\mathrm{M}}^{\mathrm{sc}\dagger} \phi^{\mathrm{sc}}$, where $H_{\mathrm{M}}^{\mathrm{sc}} = H_{\mathrm{dm}}^{\mathrm{sc}} M$.

Two key observations can now be made. Under the assumption that G_{a} is an accurate model of the NGS WFSs (i.e., neglecting nonlinear effects, system errors like misregistration, hysteresis, etc.), the NGS loop reconstructs a given DM shape into the projection of that shape onto the TT and TA modes defining the columns of M; i.e., $M\ R\ G_{\mathrm{a}}$ is a projection matrix onto the range space of M:

$$R\ G_{\mathrm{a}} = M^{\dagger} = (M^{\mathrm{T}}\ W_{\mathrm{a}}\ M)^{-1}\ M^{\mathrm{T}}\ W_{\mathrm{a}}, \quad (22)$$

$$W_{\mathrm{a}} = G_{\mathrm{a}}^{\mathrm{T}}\ W^{\mathrm{ngs}}\ G_{\mathrm{a}}, \quad (23)$$

$$M\ R\ G_{\mathrm{a}} = P_{\mathrm{a}} = M\ M^{\dagger}, \quad (24)$$

where W_{a} is a symmetric weighting matrix in actuator space. Note that the projection matrix P_{a} is W_{a} orthogonal; i.e., $P_{\mathrm{a}}^2 = P_{\mathrm{a}}$ and $P_{\mathrm{a}}^{\mathrm{T}}\ W_{\mathrm{a}} = W_{\mathrm{a}}\ P_{\mathrm{a}}$. Projecting the NGS modes out of the LGS DM actuator vector at the output of the DM fitting step (i.e., applying the operator $I - P_{\mathrm{a}}$) will fully decouple the NGS loop from the LGS loop. However, there will in practice be some amount of coupling owing to WFS nonlinearities and spatial aliasing effects. This projection does not significantly increase real-time computation requirements since its cost is on the order of the total number of actuators.

Note that in open loop, an integrated tomography solution is mandatory, as the above split approach will not provide adequate compensation of the TT/TA modes on account of spatial aliasing corrupting the open-loop NGS WFS measurements.

4. SAMPLE MONTE CARLO SIMULATION RESULTS

To test and validate the benefits of the proposed split tomography architecture, we have conducted a series of wave optics Monte Carlo simulations for the TMT LGS MCAO system using our in-house-developed Linear Adaptive Optics Simulator (LAOS) software. An overview of the TMT LGS MCAO system and its expected performance can be found in [18,19]. The reference design MCAO configuration includes order 63×63 and 75×75 wavefront correction at 800 Hz using a pair of conventional piezo-stack DMs conjugate to ranges of 0 and 12 km, six LGS WFSs of order 60×60, two near-infrared TT NGS WFSs, and one near-infrared TTFA NGS WFS. All WFSs are Shack–Hartmann wavefront gradient sensors. The main modeling assumptions were as follows: six-layer Cerro Tololo turbulence profile simulated at a resolution of 1/64 m and characterized at 500 nm and zenith by layer altitudes (relative to ground level of 2.2 km above sea level at Cerro Tololo) of 0, 2.6, 5.6, 7.7, 12.9, and 15.5 km; a Fried parameter of 15 cm; an isoplanatic angle of 2.5 arcsec; a generalized (dual-conjugate) isoplanatic angle of 10.5 arcsec; a Greenwood frequency of 29 Hz; Taylor frozen atmospheric temporal dynamics; matched reconstructed turbulence profile sampled at half the LGS WFS subaperture size, i.e., 1/4 m; telescope pointing at zenith; TMT pupil amplitude function sampled at 1/64 m, used to generate LGS WFS spots and evaluate performance; wave optics modeling of LGS WFSs using a polar coordinate detector array with 16×4 pixels per subaperture and a constrained matched-filter spot-position estimation algorithm [20]; median lidar sodium profile; an NGS asterism at the vertices of an equilateral triangle of circumradius equal to 30 arcsec at the center of a square 10 arcsec science field of view (FoV); a pair of TT NGS

WFSs and a TTFA NGS WFS modeled geometrically as RMS best-fit Zernike TT measurements (i.e., as the last two components of the amplitude-weighted pseudoinverse of the Zernike piston/tip/tilt modal matrix applied to the subaperture pupil-plane wavefront) with 1.8 milliarcsec (mas) measurement noise for the TTFA WFS and 0.4 mas for the two TT WFSs, which correspond to the expected noise for quadrant detectors imaging bright stars of J magnitude equal to 15 with a Strehl ratio of 50%, 40% end-to-end optical throughput, no read noise and a quantum efficiency of 30%; common 800 Hz sampling frequency for all WFSs; lossless integrators with gains of 1/2 and two-frame latency controlling the DMs and the tip/tilt mirror (TTM).

Sample comparative performance results for split and integrated tomography algorithms are summarized in Table 1. Relative RMS wavefront errors are quoted as quadrature differences (nm) from the closed-loop residual wavefront errors obtained with three warm-started split tomography FDPCG iterations, followed by four warm-started CG iterations for the DM fitting. Performance has been averaged over a 10 arcsec square 3×3 evaluation field with Simpson weights and 12 uncorrelated simulations of 2400 time steps each (3 s of real time each) with translating atmospheric phase screens. The total (piston-removed) error has been expressed as the root-sum-square (RSS) of errors in the LGS and NGS modes, and the error in the NGS modes has been expressed as the RSS of the TT and TA errors. For the reference algorithm, these terms are, respectively, 141 nm total, 139 nm in the LGS modes, 27 nm in the NGS modes, 17 nm in global TT, and 21 nm in the TA modes. It is seen that both integrated and split tomography provide very similar levels of closed-loop residual wavefront error in the LGS-controlled modes and that a somewhat smaller number of iterations are required by the CG and BGS-CG algorithms for the split approach. Note that the performance estimates for the NGS-controlled modes are for the case of a single, symmetric, bright asterism and that sky coverage simulations are needed to assess the relative performance of the algorithms over an ensemble of representative asterisms. Note also that imaginary numbers indicate a negative differential wavefront variance, i.e., better performance compared with the reference algorithm. Finally, the pair of empty boxes are two cases that have not been simulated since the performance in the LGS-controlled modes of the corresponding algorithms was already superior to that of the reference algorithm with fewer iterations.

5. COMPUTATION AND MEMORY REQUIREMENTS

Computation and memory requirements for the TMT LGS MCAO system are provided in Table 2 for all three algorithms for the split tomography architecture. The aperture-plane grid is 123 points wide with mesh size equal to $\Delta_0=d_{sa}/2=1/4$ m, and there are $N_{dm}=2$ DM grids with sheared actuator geometry, respectively 63 and 75 actuators wide with 1/2 m mesh size each. Requirements are given for the following most stressing atmospheric grid sizes that allow for a telescope zenith angle of up to 60 deg, and a circular science field of view of diameter of up to 120 arcsec: 6 grids 256 points wide for the FDPCG algorithm with oversampling parameter α_k equal to 1, and 6 grids respectively 123,143,161,179,110, and 122 points wide for the CG and BGS-CG algorithms with oversampling parameter equal to 1 on the four lowest grids and 2 on the remaining upper grids. The large grid size for the FDPCG algorithm arises from the fact that the algorithm performs poorly with mixed-resolution grids [15].

All of the computation and memory requirements have been obtained by computing the number of nonzero elements of a sparse matrix associated with each operator. Four bytes per real and integer have been assumed. Memory requirements for the aperture-plane operators Γ and W are listed for a sparse-matrix implementation with compressed sparse-row storage format. Finally, memory requirements for each bilinear interpolation operator have been calculated for a grid-based implementation. If $n_0=123$ denotes the aperture-plane grid size, up to n_0 integers and n_0 real numbers have to be stored for the x direction, and similarly for the y direction. There is no storage for ground-level grids (DM and atmospheric screen). The Fourier domain preconditioning matrix used by the FDPCG algorithm does not require sparse-matrix storage. This matrix is block diagonal with N/b full square blocks, where N denotes the tomography matrix system size and b the number of rows of blocks. Note that each block is Hermitian; therefore only $N(b+1)/2$ complex numbers need to be stored. The cost to apply the FD pre-

Table 1. Quadrature Difference (nm) from Closed-Loop Residual Wavefront Errors Obtained with Three Warm-Started Split Tomography FDPCG Iterations, Followed by Four Warm-Started DM Fitting CG Iterations[a]

Algorithm	No. CG Iterations	Split Tomography					Integrated Tomography				
		Total	**LGS**	NGS	TT	TA	Total	**LGS**	NGS	TT	TA
FDPCG	3	0	**0**	0	0	0	j10	**j9**	j4	5	j7
CG	30	25	**j3**	25	13	21	65	**19**	62	33	53
	40						44	**j6**	44	24	37
BGS-CG	20	23	**j3**	23	13	19	50	**13**	48	25	42
	30						30	**j12**	32	16	28

[a]The imaginary numbers (jx) indicate improved performance with respect to this baseline. Performance in the LGS modes has been highlighted in bold.

Table 2. Computation and Memory Requirements for the TMT LGS AO System for the Split Tomography Architecture[a]

Operator	Algorithm	Memory (MB)	No. CG Iterations	No. Operations (MMAC)
Total Tomography	CG	<0.8	30	178.5
	BGS-CG	<0.8	20	187.4
	FDPCG	38.3	3	191.4
Total DM Fitting	CG	<0.6	4	17
TT/TA Modes Removal		0.2		0.1

[a]Memory is quoted in units of megabytes (MB) and assumes 4 bytes per real and integer, and the number of operations is given in units of million of multiplication-accumulation computations (MMAC).

conditioning matrix has been calculated as $4Nb + 2\sum_{k=1}^{N_{ps}} n_k \log_2(n_k)$, where $n_k = 256$ and $N = \sum_k n_k^2$.

6. CONCLUSION

A split atmospheric tomography control architecture using laser and natural guide stars has been defined for MCAO. Its performance has been evaluated in extensive wave optics Monte Carlo simulations for the TMT LGS MCAO system against the delivered performance of the integrated control architecture. Three warm-started iterative algorithms have been analyzed for atmospheric tomography in both cases: a previously proposed FDPCG algorithm, a simple CG algorithm without preconditioning, and a novel layer-oriented BGS-CG algorithm. Provided that enough warm-started iterations are performed, all three algorithms yield essentially identical closed-loop residual RMS wavefront errors for both control architectures, with the caveat that a somewhat smaller number of iterations are required by the CG and BGS-CG algorithms for the split approach. These results demonstrate that the split control approach benefits from (i) a simpler formulation of minimum variance atmospheric tomography allowing for algorithms with reduced computational complexity and cost compared with FDPCG or MG, (ii) a simpler, more flexible control of the NGS-controlled modes, and (iii) a reduced coupling between the LGS- and NGS-controlled modes. Computation and memory requirements for all three algorithms have also been given for the split control approach for the TMT LGS MCAO system and appear feasible in relation to the performance specifications of current hardware technology.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the TMT partner institutions. They are the Association of Canadian Universities for Research in Astronomy (ACURA), the California Institute of Technology, and the University of California. This work was supported as well by the Gordon and Betty Moore Foundation, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada (NRC), the Natural Sciences and Engineering Research Council of Canada, the British Columbia Knowledge Development Fund, the Association of Universities for Research in Astronomy (AURA), and the U.S. National Science Foundation (NSF).

REFERENCES

1. H. Bouy, J. Kolb, E. Marchetti, E. L. Martín, N. Huélamo and D. B. Navascués, "Multi-conjugate adaptive optics images of the Trapezium cluster," Astron. Astrophys. **477**, 681–690 (2008).
2. B. L. Ellerbroek, "Efficient computation of minimum-variance wave-front reconstructors using sparse matrix techniques," J. Opt. Soc. Am. A **19**, 1803–1816 (2002).
3. B. L. Ellerbroek, L. Gilles, and C. R. Vogel, "Numerical simulations of multi-conjugate adaptive optics wavefront reconstruction on giant telescopes," Appl. Opt. **42**, 4811–4818 (2003).
4. C. R. Vogel and Q. Yang, "Fast optimal wavefront reconstruction for multi-conjugate adaptive optics using the Fourier domain preconditioned conjugate gradient algorithm," Opt. Express **14**, 7487–7498 (2006).
5. P. L. Wizinowich, D. Le Mignant, A. H. Bouchez, R. D. Campbell, J. C. Y. Chin, A. R. Contos, M. A. van Dam, S. K. Hartman, E. M. Johansson, R. E. Lafon, H. Lewis, P. J. Stromski, and D. M. Summers, "The W. M. Keck Observatory Laser Guide Star Adaptive Optics System: Overview," Publ. Astron. Soc. Pac. **118**, 297–309 (2006).
6. F. Rigaut and E. Gendron, "Laser guide star in adaptive optics: the tilt determination problem," Astron. Astrophys. **261**, 677–684 (1992).
7. B. L. Ellerbroek and G. M. Cochran, "A wave optics propagation code for multiconjugate adaptive optics," Proc. SPIE **4494**, 104–120 (2002).
8. B. L. Ellerbroek and F. Rigaut, "Methods for correcting tilt anisoplanatism in laser-guide-star-based multiconjugate adaptive optics," J. Opt. Soc. Am. A **18**, 2539–2547 (2001).
9. B. L. Ellerbroek, "Evaluating and optimizing control algorithms for combined LGS/NGS MCAO systems," Appendix E of the Conceptual Design Review Documents for MCAO for Gemini South, Gemini Observatory (2002), available online at http://www.gemini.edu/sciops/instruments/adaptive-optics/.
10. N. Kehtarnavaz and M. Gamadia, *Real-Time Image and Video Processing: From Research to Reality* (Morgan and Claypool, 2006), available online at http://www.morganclaypool.com/.
11. L. Gilles, "Closed-loop stability and performance analysis of least-squares and minimum-variance control algorithms for multi-conjugate adaptive optics," Appl. Opt. **44**, 993–1002 (2004).
12. G. A. Hyver and R. M. Blankinship, "ALI high power beam control," Adv. Astronaut. Sci. **88**, 445–460 (1995).
13. B. L. Ellerbroek and C. R. Vogel, "Simulations of closed-loop wavefront reconstruction for multi-conjugate adaptive optics on giant telescopes," Proc. SPIE **5169**, 206–217 (2003).
14. L. H. Lawton, "Loopshaped wavefront control using open-loop reconstructors," Opt. Express **14**, 7477–7486 (2006).
15. L. Gilles, B. L. Ellerbroek, and C. R. Vogel, "A comparison of multigrid V-cycle versus Fourier domain preconditioning for laser guide star atmospheric tomography," in *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis*, Optical Society of America Topical Meeting Technical Digest 2007, paper JTuA1, available online at http://www.opticsinfobase.org.
16. L. Lessard, M. West, D. MacMynowski, and S. Lall, "Warm-started wavefront reconstruction for adaptive optics," J. Opt. Soc. Am. A **25**, 1147–1155 (2008).

17. Y. Saad, *Iterative Methods for Sparse Linear Sytems*, 2nd ed. (SIAM, 2003).
18. B. Ellerbroek, S. Adkins, D. Andersen, J. Atwood, C. Boyer, P. Byrnes, R. Conan, L. Gilles, G. Herriot, P. Hickson, E. Hileman, D. Joyce, B. Leckie, M. Liang, T. Pfrommer, J.-C. Sinquin, J.-P. Véran, L. Wang, and P. Welle, "Progress towards developing the TMT adaptive optical systems and their components," Proc. SPIE **7015** (2008) (to be published).
19. L. Gilles, L. Wang, and B. Ellerbroek, "Wavefront error budget development for the Thirty Meter Telescope laser guide adaptive optics system," in "Adaptive Optics Systems," Proc. SPIE **7015**, (2008) (to be published).
20. L. Gilles and B. L. Ellerbroek, "Constrained matched filtering for extended dynamic range and improved noise rejection for Shack–Hartmann wavefront sensing," Opt. Lett. **33**, 1159–1161 (2008).