Computationally Efficient Wavefront Reconstruction for Multi-Conjugate Adaptive Optics (MCAO)

Brent Ellerbroek

AURA New Initiatives Office

IPAM Workshop on Estimation and Control Problems in Adaptive Optics

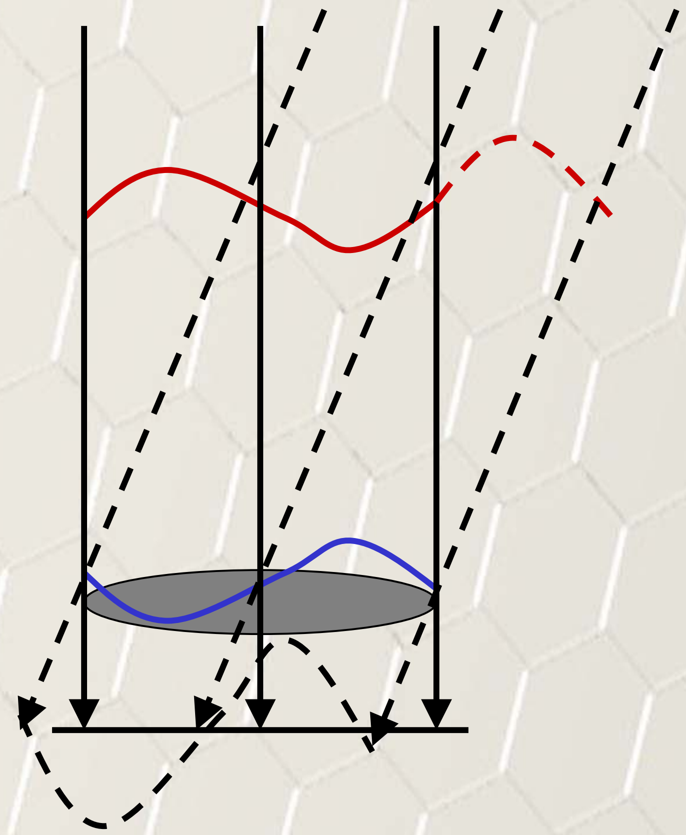
January 23, 2004

Presentation Outline

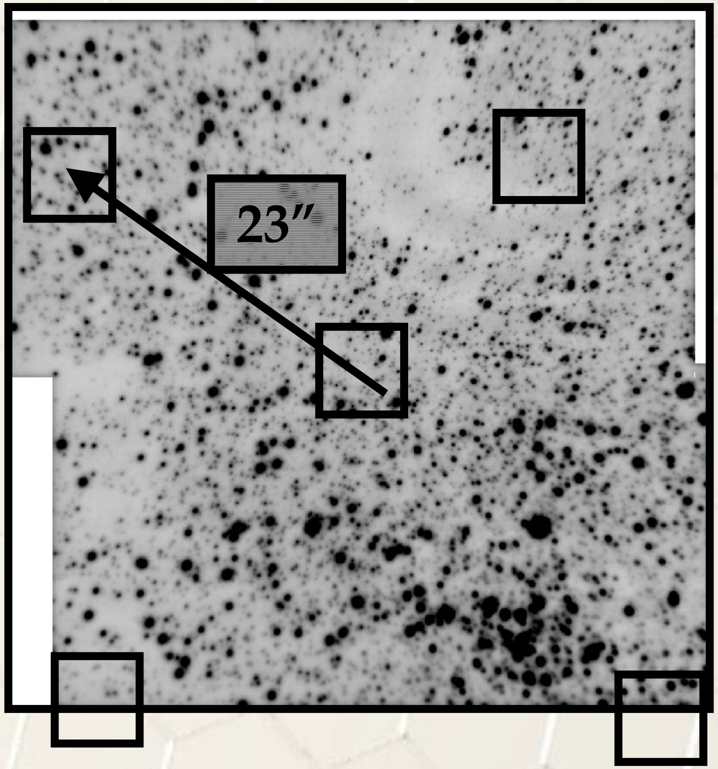
- Anisoplanatism and MCAO
- Minimum variance wavefront reconstruction methods for MCAO
  - Formulation, analytical solution, and scaling issues
- Computationally efficient methods for very high-order MCAO systems
  - Spatial frequency domain modeling (Tokovinin)
  - Sparse matrix techniques
  - Conjugate gradients with multigrid preconditioning
- Sample simulation results
  - MCAO Performance scaling with telescope diameter
- Summary, acknowledgements, references

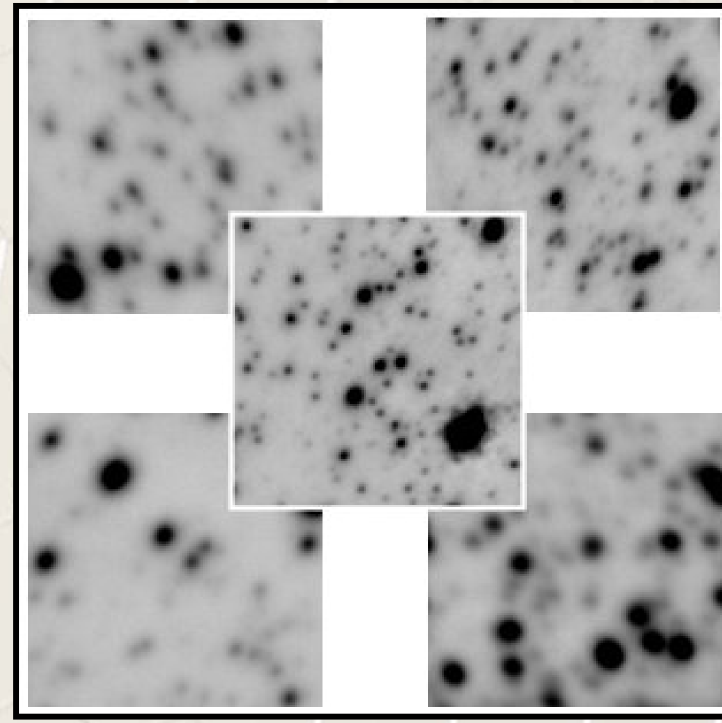
Anisoplanatism and Adaptive Optics

- Bright guidestars are needed for wavefront sensing
  - Not enough bright natural stars for astronomical applications
  - Progress is being made in using lasers to generate artificial stars
- Even with lasers, the corrected field-of-view is limited
  - Turbulence is 3-dimensional
  - One deformable mirror provides correction in a single direction
  - Anisoplanatism



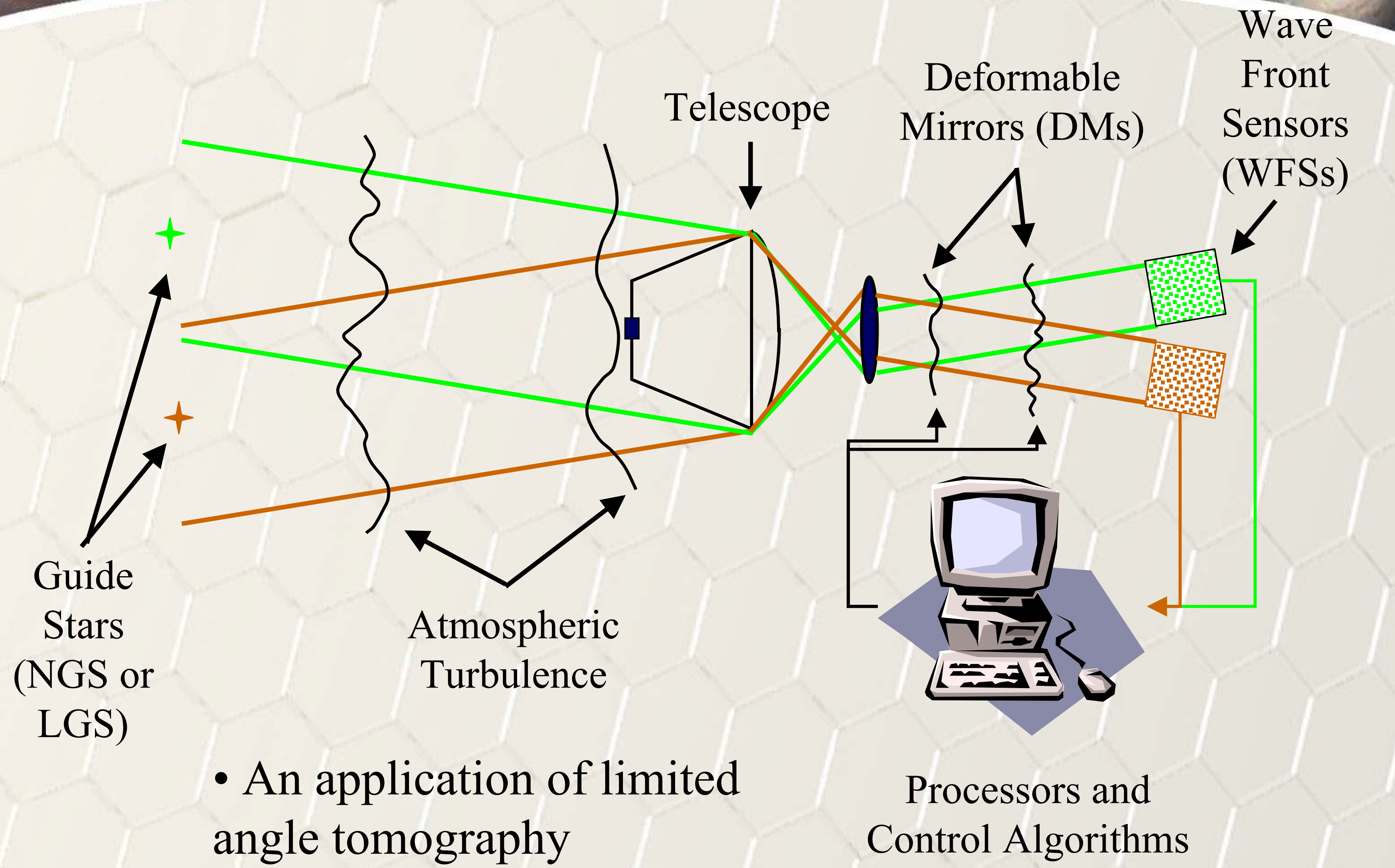
Adaptive Optics Imagery with Anisoplanatism





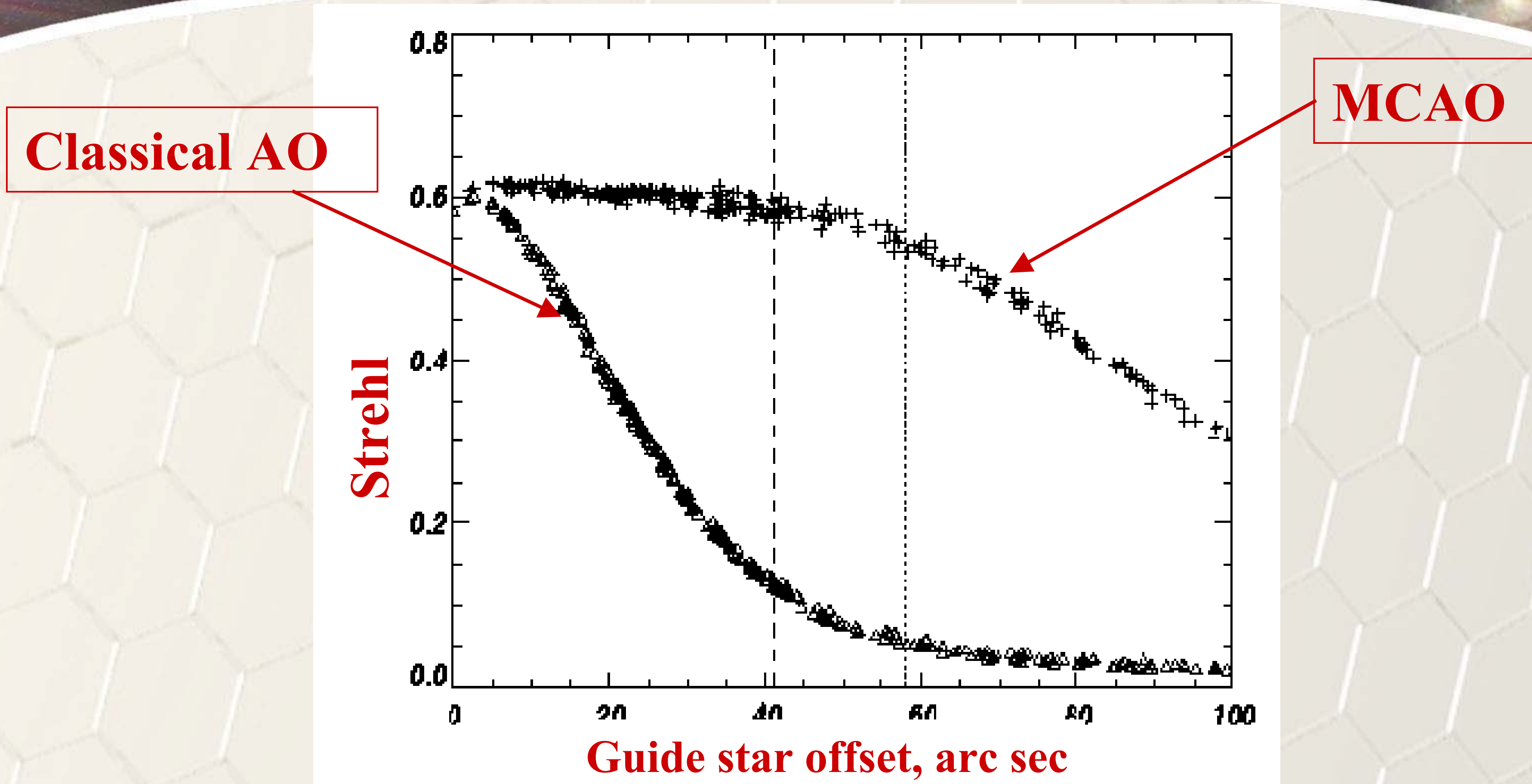
- Low-order AO system on the Gemini-North telescope
- Ambient seeing: 0.9”
- AO-compensated seeing: 0.12” (center of field) to 0.19” (corner)
- Impact increases as the quality of correction improves

MCAO Compensates Turbulence in Three Dimensions



- An application of limited angle tomography

Sample MCAO Simulation Result



- 8 meter telescope
- 2 deformable mirrors with 13 by 13 actuators
- 5 wavefront sensors with 12 by 12 subapertures

Wavefront Reconstruction for MCAO is Challenging

- Multiple turbulence layers, deformable mirrors, wavefront sensors
- Richer cross-coupling between variables
- Higher dimensionality estimation problem
  - Especially for future extremely large telescopes!
- Wide-field performance evaluation and optimization

Wavefront Reconstruction as a Linear Inverse Problem

- Quantities of interest
  - Turbulence profile x…
  - …to be corrected by a DM actuator command vector a…
  - … using a WFS measurement s with noise component n…
  - …leaving a residual phase error ϕ with mean-square value σ^2
- Relationships
  - $s = Gx + n$ (wavefront sensing)
  - $a = Rs$ (wavefront reconstruction)
  - $\phi = H_x x - H_a a$ (residual error computation)
  - $\sigma^2 = \phi^T W \phi$ (variance evaluation)
- Objective: Select R to (in some sense) minimize σ^2

Minimum Variance Wavefront Reconstruction

- Model x, s, and n as zero mean random variables with finite second moments
- Select R to minimize $\langle \sigma^2 \rangle$ (the expected value of σ^2):

$$R_* = \arg\min_R \langle \sigma^2 \rangle$$

$$= \arg\min_R \left\langle \left[H_x x - H_a R s \right]^T W \left[H_x x - H_a R s \right] \right\rangle$$

- Partials of $\langle \sigma^2 \rangle$ with respect to R_{ij} must vanish at $R=R_*$
- Solution given by $R_*=F_*E_*$, where

$$E_* = \langle xs^T \rangle \langle ss^T \rangle^{-1} = \left(G^T \langle nn^T \rangle^{-1} G + \langle xx^T \rangle^{-1} \right) G^T \langle nn^T \rangle^{-1}$$

$$F_* = \left(H_a^T W H_a \right)^{-1} H_x^T W H_x$$

Interpretation (and Use) of $R_* = F_* E_*$

- Interpretation
  - E_* is the "turbulence **E**stimation matrix"
    - Minimum variance estimate of profile x from measurement s
    - Depends upon WFS geometry, statistics of x and s
    - Independent of the DM geometry
  - F^* is the "turbulence **F**itting matrix"
    - RMS best fit to estimated value of x using DM degrees of freedom
    - Independent of WFS geometry, statistics of x and s
    - Depends upon the DM geometry
- Use
  - Once R_* is known, we can estimate performance using

$$\min_R \langle \sigma^2 \rangle = \langle [H_x x - H_a R_* s]^T W [H_x x - H_a R_* s] \rangle$$

  - …or we can use R_* to run simulations (or even systems)

Computational Complexity

- R_* has complexity $O(N^3)$ to explicitly compute and evaluate, complexity $O(N^2)$ to apply in real time
  - Must be computed/evaluated in a few hours for studies
  - Must be applied at rates of 1-2 KHz for actual use
- Current generation MCAO systems have $N < 1000$
  - Computationally feasible
- Proposed MCAO systems have $N > 10^4$ or 10^5
  - Explicit computations inefficient or outright infeasible
  - How do we analyze and simulate such systems???

Analytical Methods in the Spatial Frequency Domain

- Wavefront propagation, sensing, correction, and reconstruction are all approximately spatial filtering operations
- Filtering representation becomes exact in the limit of an infinite aperture AO system
- Wavefront reconstruction decouples into small independent problems at each spatial frequency
  - Each problem has dimensionality $2\ N_{wfs}$ by N_{dm}
- Overall complexity scales as $O(N_{freq})\ \alpha\ O(N)$
- Analytical method only, but very useful

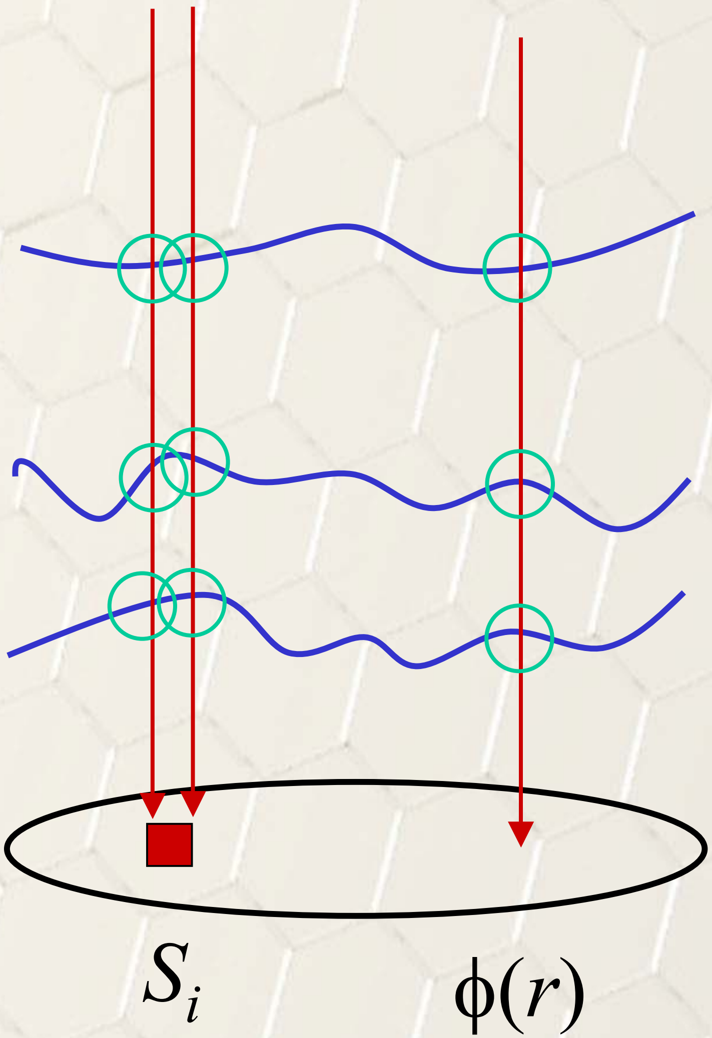
Efficient Approaches for the Spatial Domain

- Must solve $Ax=y$, where

  $$A = G^T \langle nn^T \rangle^{-1} G + \langle xx^T \rangle^{-1} \quad \text{or} \quad A = H_a^T W H_a$$

  without explicitly computing A^{-1}
- Exploit matrix structure
  - G, H_a, W are sparse
  - $\langle nn^T \rangle$ is diagonal (plus a low-rank perturbation due to laser guide star position uncertainty)
  - $\langle xx^T \rangle^{-1}$ has good approximations that are sparse
- Efficient solutions possible
  - Sparse matrix techniques (close, but not quite)
  - Conjugate gradients with multigrid preconditioning

G, H Sparse for Nodal Representations of Turbulence



- Each value of $\phi(r)$ is determined by turbulence values along a single ray path
- Each WFS measurement s_i is determined by values of $\phi(r)$ within a small subaperure

Sparse Matrix Methods

- Suppose A is sparse (with bandwidth $O(N^{1/2})$)
- Factor

  $A = LL^T$

  where L is sparse and lower triangular
- Solve $Ax=y$ in two steps:

  $Lx' = y$, followed by $L^Tx = x'$
- Complexity reduced from $O(N^2)$ to $O(N^{3/2})$
- Complexity further reduced by reording rows/columns of A
- For F_*, $A = H_a^T W H_a$ is sparse (at least for conventional AO)
- For E_*, $A = G^T \langle nn^T \rangle^{-1} G + \langle xx^T \rangle^{-1}$ isn't sparse for two reasons:
  - The turbulence covariance matrix $\langle xx^T \rangle$ isn't sparse
  - For laser guidestars, $\langle nn^T \rangle$ is the sum of sparse and low rank terms

Sparse Approximation to Turbulence Statistics

- $\langle xx^T \rangle^{-1}$ is block diagonal, with N_{layer} by N_{layer} blocks
  - Each diagonal block is full rank!
- We approximate block j as $\alpha_i^{-1} D^T D$
  - α_i proportional to layer strength
  - D is a discrete (and sparse) approximation to ∇^2
- Heuristic justification #1:
  - Both $\langle xx^T \rangle^{-1}$ and $D^T D$ suppress high spatial frequencies
- Heuristic justification #2:
  - In the spatial frequency domain

$$\left\langle \hat{x}(\kappa)\hat{x}^*(\kappa') \right\rangle \propto \delta(\kappa-\kappa')\kappa^{-11/3} \approx \delta(\kappa-\kappa')\kappa^{-4}$$

$$\left\langle \hat{x}(\kappa)\hat{x}^*(\kappa) \right\rangle^{-1} \propto \kappa^4 = \kappa^2\kappa^2 \propto \left[\mathrm{FT}\left(\nabla^2\right)\right]^T \left[\mathrm{FT}\left(\nabla^2\right)\right]$$

LGS Measurement Noise

- For a LGS WFS, *n* is determined by two effects:
  - Detector readout noise and photon statistics (uncorrelated)
  - LGS position uncertainty on the sky
    - Two dimensions of uncertainty per guidestar, correlated between subapertures
- More formally

$$n = n_r + n_t$$

$$\left\langle nn^T \right\rangle = \left\langle n_r n_r^T \right\rangle + \left\langle n_t n_t^T \right\rangle = \mathrm{diag}(\sigma_i^2) + \sigma_t^2 UU^T$$

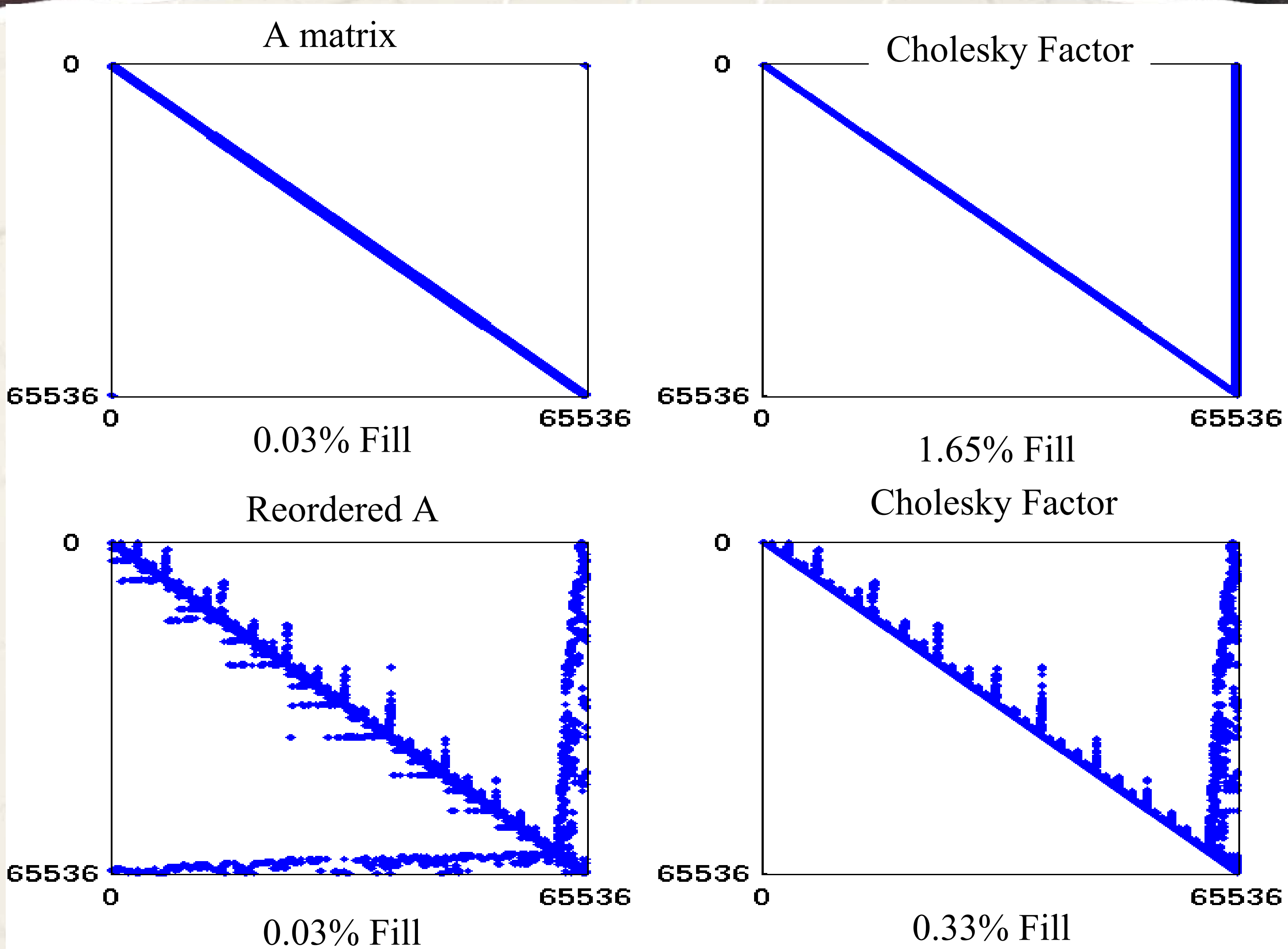
  - UU^T is a non-sparse matrix of rank 2 N_{LGS}
- Sparse matrix methods are not immediately applicable

Applying the Matrix Inversion Lemma

$$\left(M - UV^T\right)^{-1} = M^{-1} + \left(M^{-1}U\right)\left(I - V^T M^{-1} U\right)^{-1}\left(M^{-1}V\right)^T$$

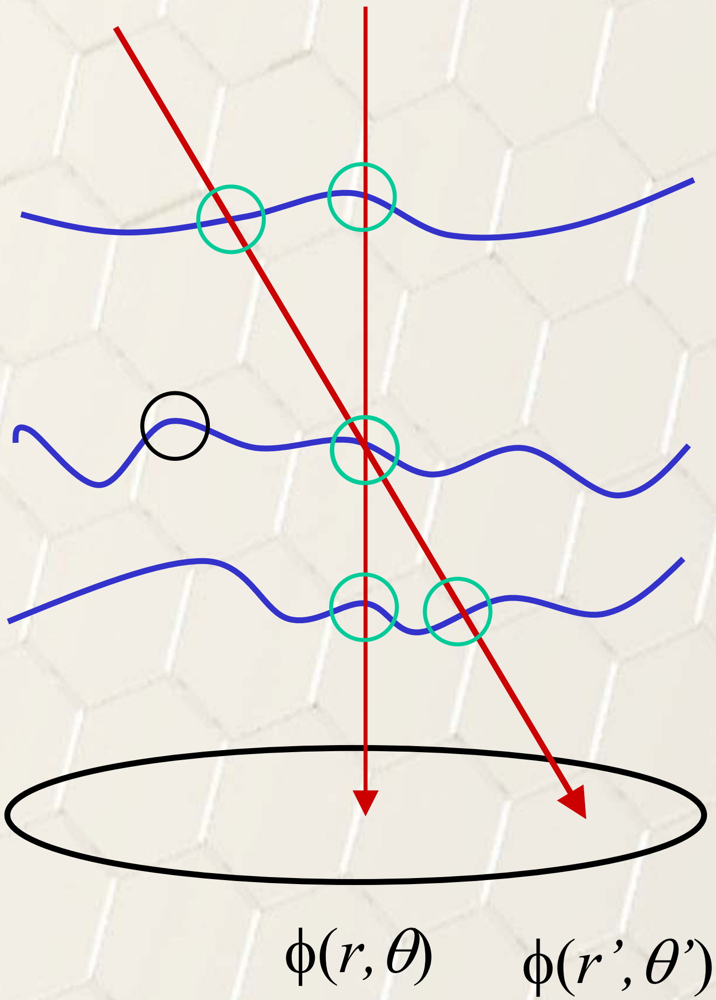
- $\left\langle nn^T \right\rangle^{-1} = \left\langle \mathrm{diag}\left(\sigma_i^2\right) + \sigma_T^2 UU^T \right\rangle^{-1}$ is the sum of $\mathrm{diag}\left(\sigma_\mathrm{i}^{-2}\right)$ and a low rank term UU^T
- $$G^T \left\langle nn^T \right\rangle^{-1} G + \left\langle xx^T \right\rangle^{-1} = \underbrace{G^T \mathrm{diag}\left(\sigma_\mathrm{i}^{-2}\right) G + \left\langle xx^T \right\rangle^{-1}}_{\text{Sparse}} + \underbrace{\left(GU'\right)\left(GU'\right)^T}_{\text{Low Rank}}$$
- Can solve $\left(G^T \left\langle nn^T \right\rangle^T G + \left\langle xx^T \right\rangle^{-1}\right)^{-1} x = y$
  by solving $\left(G^T \mathrm{diag}\left(\sigma_i^2\right) G + \left\langle xx^T \right\rangle^{-1}\right)^{-1} x = y$
  and adding a perturbation term depending upon
  $\left(G^T \mathrm{diag}\left(\sigma_i^2\right) G + \left\langle xx^T \right\rangle^{-1}\right)^{-1} (GU')$

Sample Matrix Factorizations for E_*



- Conventional AO with 1 DM and 1 WFS!

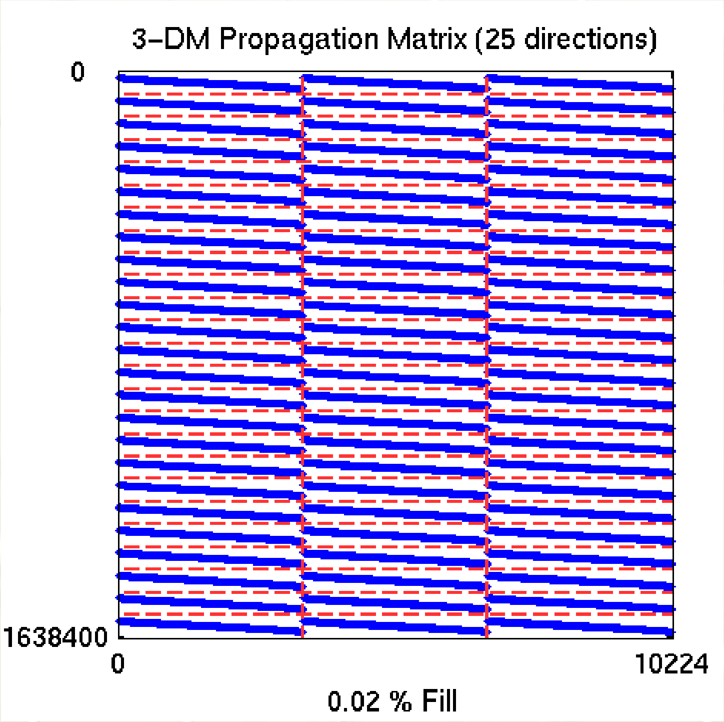
MCAO Increases Coupling between Turbulence Layers



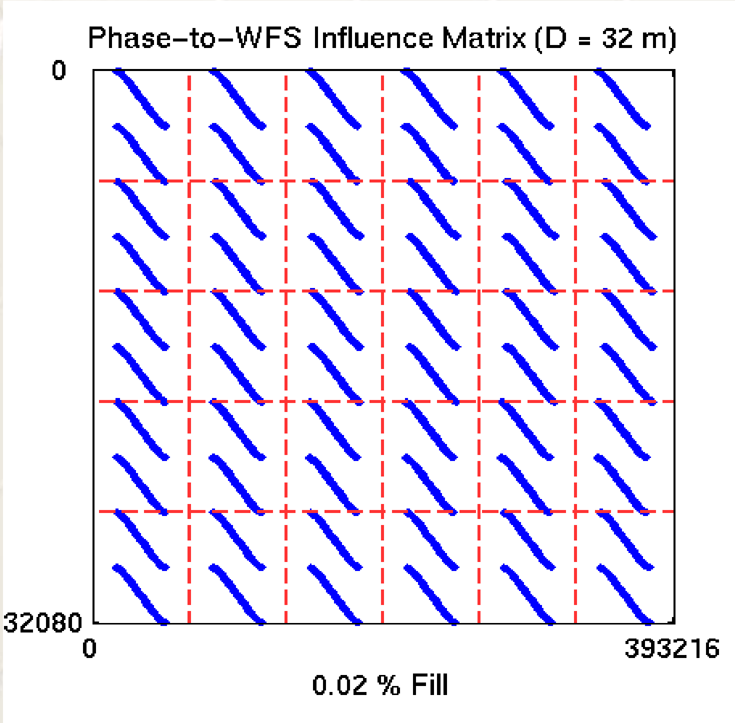
- However, the coupling within a single layer is no greater than before

G,H Matrices Are Block Structured for MCAO

- Column block structure due to multiple atmospheric layers
- Row block structure due to multiple stars/guidestars

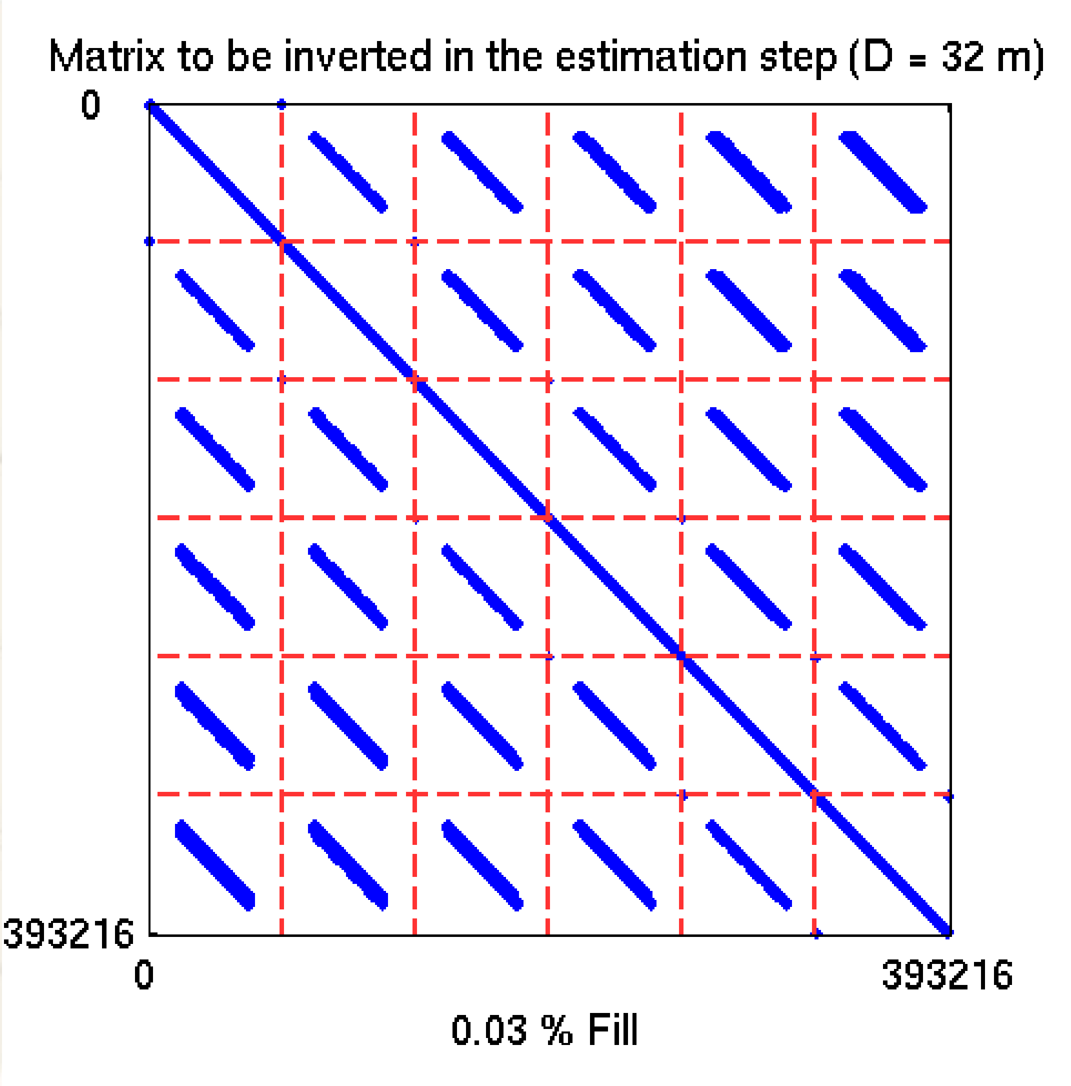


H_a (3 mirrors, 25 stars)



G (5 guidestars, 6 atmospheric layers)

Cross-Coupling of Atmospheric Layers for MCAO



- Fill-in of "sparse" Cholesky factorization exceeds 10%
- Cannot factor matrices for a 32m diameter system in a 2 Gbyte address space

An "Efficient" MCAO Reconstruction Algorithm

- Biggest challenge is solving $Ax=y$ with

$$A = G^T \left\langle nn^T \right\rangle^{-1} G + \left\langle \delta\delta^T \right\rangle^{-1}$$

- Minimize $||Ax-y||^2$ using conjugate gradients
- Use multigrid preconditioning to accelerate convergence
  - Preconditioning: Solve an approximate system $A'x=y$ once per conjugate gradient cycle
  - Multigrid: Solution to $A'x=y$ determined on multiple spatial scales to accelerate convergence at all spatial frequencies
- Solution on each multigrid scale is determined using a customized (new?) technique:
  - Block symmetric Gauss-Seidel iterations on $Ax=y$
  - Block structure derived from atmospheric layers
  - Sparse matrix factorization of diagonal blocks

Block Symmetric Gauss-Seidel Iterations

- Blocks of A, x, y denoted as A_{ij}, x_{i}, y_{j}
- Decompose

  $A = L + D + U$

  into a sum of lower triangular, diagonal, and upper triangular blocks
- Iterative solution to $Ax = (L+D+U)x = y$ given by

$$(L+D)x'(n) = y - Ux(n)$$

$$(U+D)x(n+1) = y - Lx'(n)$$

- Solve for $x'(n)$ and $x(n+1)$ one block at a time:

$$D_i x_i'(n) = y_i - \sum_{j>i} A_{ij} x_j(n) - \sum_{j<i} A_{ij} x_j'(n)$$

$$D_i x_i(n+1) = y_i - \sum_{j<i} A_{ij} x_j'(n) - \sum_{j>i} A_{ij} x_j(n+1)$$

- Solve systems $D_i u = v$ using sparse Cholesky factorizations

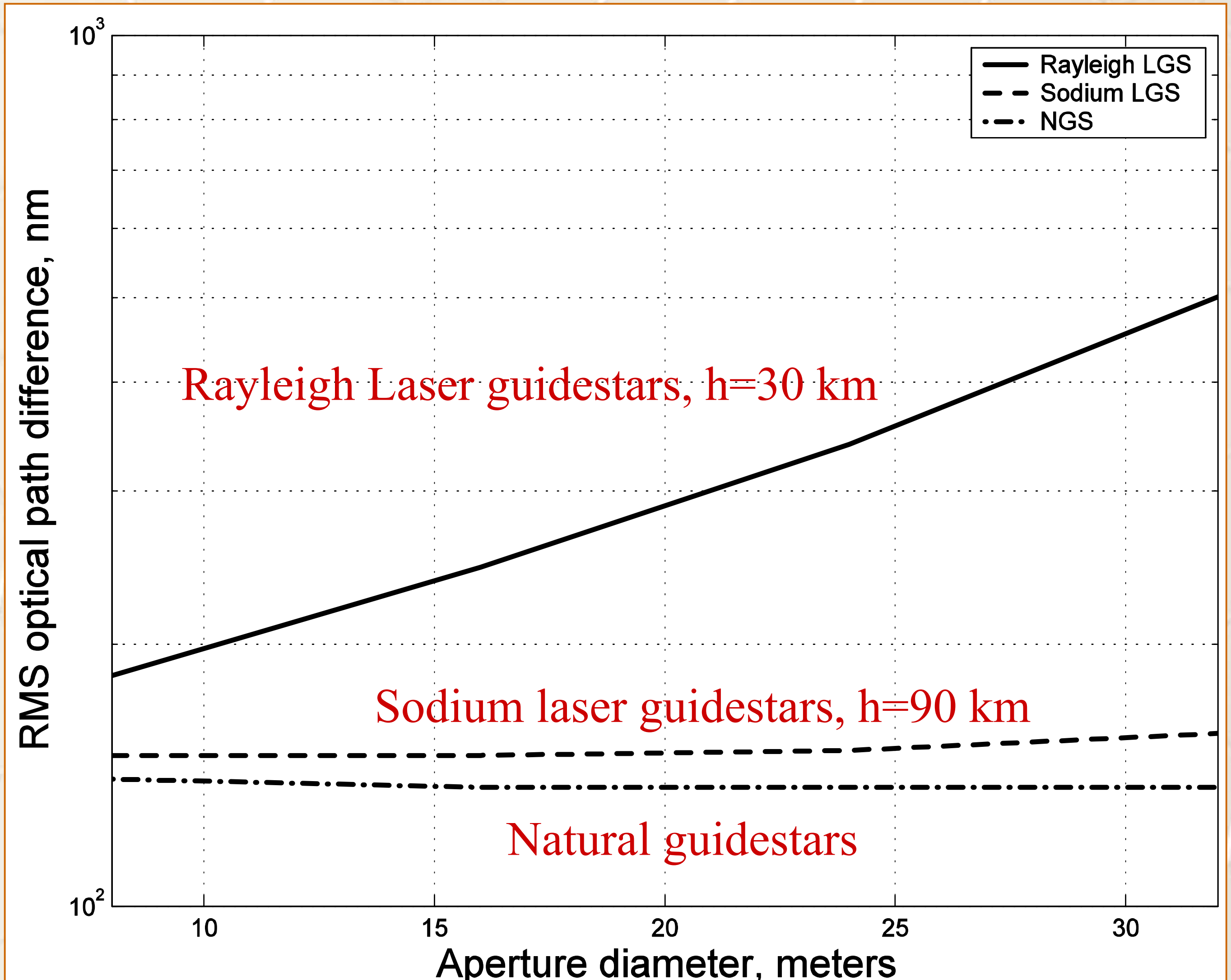
MCAO Simulations for Future Telescopes

- Goal: Evaluate MCAO performance scaling with aperture diameter *D* from *D*=8m to *D*=32m
- Consider Natural, Sodium, and Rayleigh guidestars
- Other simulation parameters:
  - Cerro Pachon turbulence profile with 6 layers
  - 1 arc minute square field-of-view
  - 3 DM's conjugate to 0, 5.15, and 10.30 km
    - Actuator pitches of 0.5, 0.5, and 1.0 m
  - 5 higher order guidestars at corners and center of 1' field
    - 0.5 m subapertures
  - 4 tip/tilt NGS WFS for laser guide star cases
  - 10 simulation trials per case using 64 m turbulence screens with 1/32m pitch

Simulation Dimensionality

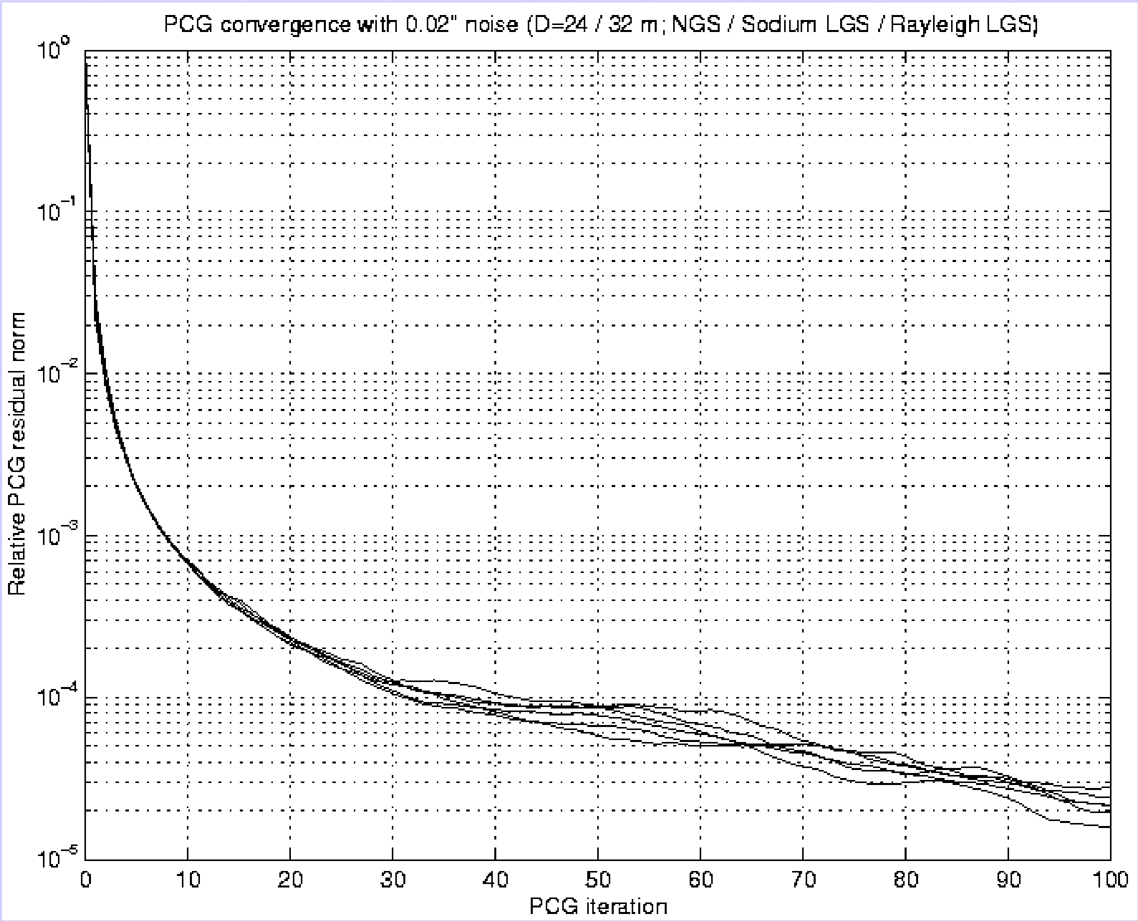
Aperture, m	8	16	24	32
WFS measurements	2240	8560	18840	33320
Phase points estimated (E_*)	7270	21226	42334	70838
DM actuators fit (F_*)	789	2417	4957	8449

Sample Numerical Results



CG Convergence Histories

- Rapid convergence for first 20 iterations
- Convergence then slows due to poor conditioning of A
- Not an issue for practical simulations
- Results effectively independent of aperture diameter and guide star type



Summary

- MCAO compensates anisoplanatism and corrects for the effects of atmospheric turbulence across extended fields-of-view
- Minimum variance estimation is a viable approach to MCAO wavefront reconstruction
- Computationally efficient methods needed for the very high order systems proposed for future extremely large telescopes
- Conjugate gradient wavefront reconstruction using multigrid preconditioning and block symmetric Gauss-Seidel iterations enables simulations of 32 meter MCAO systems with 30k sensor measurements and 8k mirror actuators
- Challenging problems remain
  - Closed-loop wavefront reconstruction and control
  - Hardware and software for real-time implementation

Acknowledgements

- Luc Gilles and Curt Vogel
  - Ongoing collaboration on efficient methods
  - Matrix sparsity plots
- Francois Rigaut
  - MCAO figure and performance plot
- Gemini Observatory
  - Sample AO results
- Support from AFOSR, NSF, and CfAO

References

- Adaptive optics websites
  - CfAO, http://cfao.ucolick.org
  - Gemini AO web pages at http://www.gemini.edu
- Minimum variance wavefront reconstruction
  - Wallner, JOSA 73, 1771 (1983)
  - Ellerbroek, JOSA A 11, 783 (1994)
  - Fusco et al., JOSA A 18, 2527 (2001)
- Efficient implementations
  - Ellerbroek, JOSA A 19, 1803 (2002)
  - Ellerbroek, Gilles, Vogel, SPIE Proc. 4839, 989 (2002)
  - Gilles, Ellerbroek, Vogel, Appl. Opt. 42, 5233 (2003)