Sparse Matrix Techniques for MCAO

Luc Gilles `lgilles@mtu.edu`

Michigan Technological University, ECE Department

Brent Ellerbroek `bellerbroek@gemini.edu`

Gemini Observatory

Curt Vogel `vogel@math.montana.edu`

Montana State University, Math Sciences

November 25, 2002

Wavefront Estimation from Idealized WFS Data

- Open-loop WFS: $s = \mathcal{G}\varphi + \eta$
- η and φ random with known statistics $C_\eta = \left\langle \eta\, \eta^T \right\rangle$ and $C_\varphi = \left\langle \varphi\, \varphi^T \right\rangle$

  $\Rightarrow \hat{\varphi} = \mathcal{G}^\dagger s$ (noise-weighted pseudo-inverse)

$$\hat{\varphi} = \text{arg} \min_{\varphi} J(\varphi)$$

$$J(\varphi) = \frac{1}{2}\left[\|\eta\|^2_{C_\eta^{-1}} + \|\varphi\|^2_{C_\varphi^{-1}}\right] = \frac{1}{2}\left[\left\langle \eta\,, C_\eta^{-1}\eta \right\rangle + \left\langle \varphi\,, C_\varphi^{-1}\varphi \right\rangle\right]$$

$$\Rightarrow \hat{\varphi} = \mathcal{G}^\dagger s \;\; \text{with} \;\; \mathcal{G}^\dagger = \left(\mathcal{G}^T C_\eta^{-1}\, \mathcal{G} + C_\varphi^{-1}\right)^{-1} \mathcal{G}^T C_\eta^{-1}$$

- Block-layered structure for MCAO
- $(\text{SNR})^2$ = $\left\langle \|\mathcal{G}\varphi\|^2 \right\rangle / \left\langle \|\eta\|^2 \right\rangle$ = $\left\langle \|\mathcal{G}\varphi\|^2 \right\rangle / (2\, n_z\, \sigma^2)$

Key Approximations to Inverse Turbulence Covariance

Approximation $\#1$

- C_φ is block diagonal (different layers are statistically independent)
- For each turbulent layer l, C_{φ_l} is BTTB (phase structure function)
- Approximate BBTB matrix by BCCB matrix (diagonalized by DFT)
- $C_{\varphi_l} \simeq$ BCCB$(\tau_l) = \mathcal{F}^{-1}$diag$(\hat{\tau}_l)\mathcal{F}$ where $\hat{\tau}_l = \mathcal{F}\tau_l$ (eigenvalues)
- $\hat{\tau}_l = c_l\,|\kappa|^{-11/3}$ (Kolmogorov PSD)

$$\Rightarrow C_{\varphi_l}^{-1} \simeq \text{BCCB}(\Lambda_l) = \mathcal{F}^{-1}\,\text{diag}(\lambda_l)\,\mathcal{F}\,, \quad \lambda_l = \hat{\tau}_l^{-1} = c_l^{-1}\,|\kappa|^{11/3}$$

Key Approximations to Inverse Turbulence Covariance

Approximation $\#2$

- BCCB(Λ_l) is a full matrix.
- Entries Λ_l rapidly decay to small values
- Approximate BCCB(Λ_l) by $\mathcal{S}_l^2$ where $\mathcal{S}_l = c_l' \, \mathcal{S}$ and $\mathcal{S}$ is the discrete Laplacian matrix with periodic boundary conditions

$$\mathcal{S} = \text{BCCB}(v) = \mathcal{F}^{-1}\text{diag}(\hat{v})\mathcal{F} \quad \text{(sparse BCCB)}$$

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

- $\mathcal{S} = \mathcal{S}_y \otimes I_x + I_y \otimes \mathcal{S}_x$ where $\mathcal{S}_y$ and $\mathcal{S}_x$ are 1D versions

Key Approximations to Inverse Turbulence Covariance

Sparsity Patterns in 1D

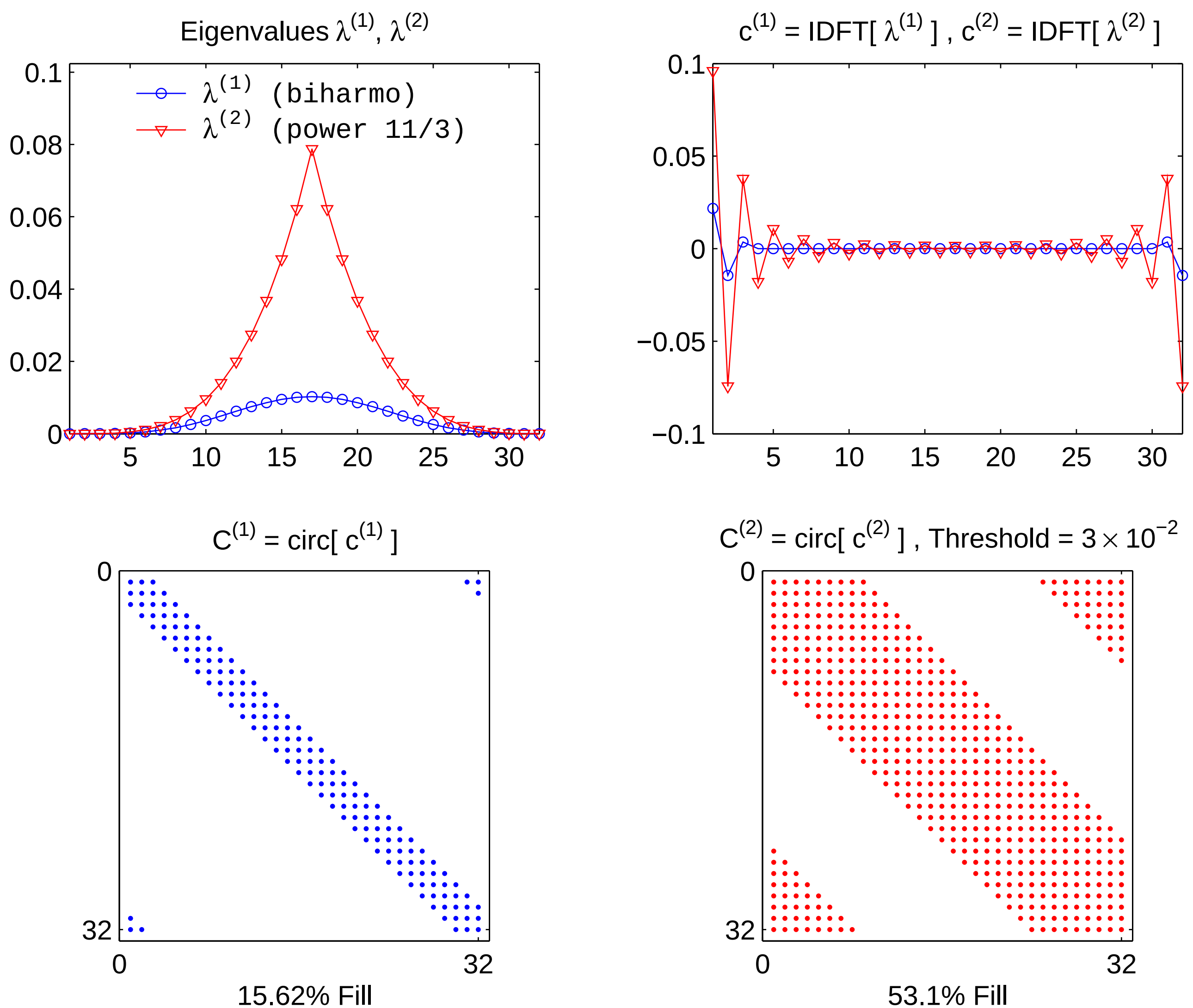
$$[\mathcal{S}\,u]_i = \left(-u_{i-1} + 2u_i - u_{i+1}\right)/\Delta x^2$$

$$\mathcal{S} = \text{circ}(v)\,, \quad v = (2,-1,0,\cdots,0,-1)^T/\Delta x^2 = (2\,e_0 - e_1 - e_{n-1})/\Delta x^2$$

$$\Rightarrow \mathcal{S} = \mathcal{F}^{-1}\text{diag}(\hat{v})\mathcal{F}\,, \quad \hat{v} = \mathcal{F}v = (2\,\hat{e}_0 - \hat{e}_1 - \hat{e}_{n-1})/\Delta x^2$$

$$\hat{v} = 4\sin^2(\pi\kappa\Delta x)/\Delta x^2 \ \ \text{(eigenvalues)} \quad \underset{\Delta x \to 0}{\longrightarrow} 4\pi^2\kappa^2$$

Key Approximations to Inverse Turbulence Covariance



MCAO Minimum Variance Reconstructor

Problem: find optimal actuator commands $\hat{a}$ minimizing MCAO wide-field error metric $\mathcal{W}$

$$\hat{a} = \mathcal{R} s$$

$$\mathcal{R} = \text{arg } \min_{R} J(R)$$

$$J(R) = \left\langle \, \|\epsilon\|^2_{\mathcal{W}} \, \right\rangle = \left\langle \, \epsilon^T \mathcal{W} \epsilon \, \right\rangle$$

$$\epsilon = H^a \, \hat{a} - H^{\varphi} \, \varphi \quad \text{(aperture-plane residual phase)}$$

$$\Rightarrow \mathcal{R} = \underbrace{H^{a\,\dagger}}_{Fitting} \; H^{\varphi} \; \underbrace{\mathcal{G}^{\dagger}}_{Estimation}$$

$$\mathcal{G}^{\dagger} = (\underbrace{\mathcal{G}^T \, C_{\eta}^{-1} \, \mathcal{G}}_{Sparse+low-rank\ (LGS)} + \underbrace{C_{\varphi}^{-1}}_{Sparse\ approx})^{-1} \, \mathcal{G}^T C_{\eta}^{-1}$$

$$H^{a\,\dagger} = \left(H^{a\,T} \, \mathcal{W} \, H^a + \text{low-rank} \right)^{-1} \; H^{a\,T} \, \mathcal{W}$$

Multigrid (MG) Methods

- Can sometimes be used as "stand-alone" system solvers.
- Can be used as preconditioners.
- Rely on multiple scales (grid sizes) inherent in certain problems.
  - Need "smoother" which damps out high-frequency components of error on fine grids.
    - Classical Gauss-Seidel iteration works well for Laplace's equation.
  - Remaining low frequency error is well-represented on coarser grids.
- Are recursive versions of the following 2-grid scheme.

2-Grid Scheme

$x_h \leftarrow S(x_h, y_h, ...)$

$r_h \leftarrow A_h x_h - y_h$

Restrict $r_H \leftarrow I_h^H r_h$

Solve $A_H e_H = r_H$

Interpolate $e_h \leftarrow I_H^h e_H$

$x_h \leftarrow x_h + e_h$

$x_h \leftarrow S(x_h, y_h, ...)$

- $S(v, w, ...)$ denotes application of smoother to solve $Ax = w$ with initial guess $x = v$.
- To obtain MG V-cycle, apply 2-grid scheme recursively. Carry out Solve step with (e_H, r_H) in place of (x_h, y_h).

Multigrid (MG) Methods

- Inter-grid transfers (restriction, or up-binning, and interpolation, or down-binning) are cheap.
- Cost is typically dominated by smoother application on finest grid.
- Choice of smoother is problem-dependent.
  - Block (i.e., layer-oriented) symmetric Gauss-Seidel (B-SGS) works well for MCAO estimation step.
  - FFT-based modified Richardson iteration works well for Ex-AO estimation.

Block Gauss-Seidel Smoother

Based on block $L + D + U$ splitting

$$A = \underbrace{\begin{bmatrix} 0 & 0 & \dots & 0 \\ A_{21} & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ A_{n1} & \dots & A_{n,n-1} & 0 \end{bmatrix}}_{L} + \underbrace{\begin{bmatrix} A_{11} & 0 & \dots & 0 \\ 0 & A_{22} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_{nn} \end{bmatrix}}_{D}$$

$$+ \underbrace{\begin{bmatrix} 0 & A_{12} & \dots & A_{1n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{n-1,n} \\ 0 & \dots & 0 & 0 \end{bmatrix}}_{U}$$

Block Gauss-Seidel Smoother

$Ax = b$ is equivalent to $(L + D)x = b - Ux$. This motivates the block forward iteration

$$(L + D)x_{k+1} = b - Ux_k, \;\; k = 0, 1, \ldots .$$

Similarly, we obtain the block backward interation

$$(D + U)x_{k+1} = b - Lx_k, \;\; k = 0, 1, \ldots .$$

Block symmetric Gauss-Seidel (B-SGS) is obtained by interweaving forward and backward iterations.

Efficient PCG Solver

Estimation Step

- MG preconditioned CG
- Block SGS smoother: requires only inversion of diagonal blocks. Implemented using reordering + full Cholesky factorization.

Fitting Step

- Incomplete Cholesky preconditioned CG. Incomplete Cholesky applied to full sparse matrix without reordering.

Preliminary Results

- ⋆ Algorithm tested against conventional matrix multiply reconstructors on 8m class problems with $\sim 7\,10^3$ degrees of freedom
- ⋆ 32m class problems with $\sim 7\,10^4$ degrees of freedom solvable in Matlab with 2-3Gb memory
- ⋆ Convergence obtained in 2-10 iterations
  - Convergence rate a strong function of WFS noise level
  - Weak function of problem dimensionality, NGS vs. LGS MCAO

Sample MCAO Problem Dimensionality

Aperture diameter (m)	8	16	32
WFS measurements	2240	8560	33320
Turbulence phase points	7270	21226	70838
DM actuators	789	2417	8449

Sample MCAO Problem Dimensionality

Estimation matrix to be inverted. 6 layers, 5 NGS's. D=16m. Phase screens 128×128.

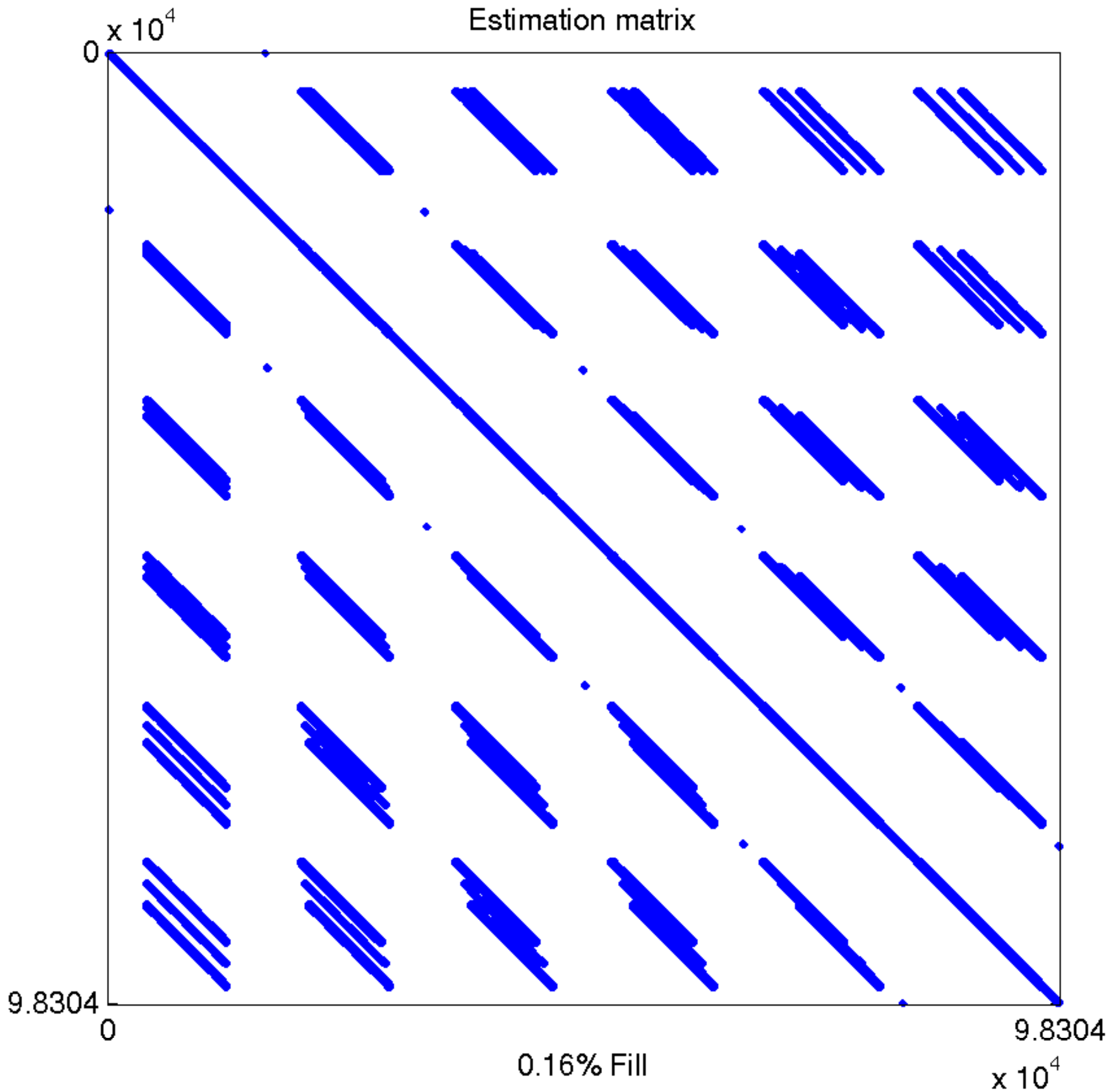


Fig.1

Top layer diagonal block of estimation matrix. D=16m. Phase screens 128 $\times$ 128.

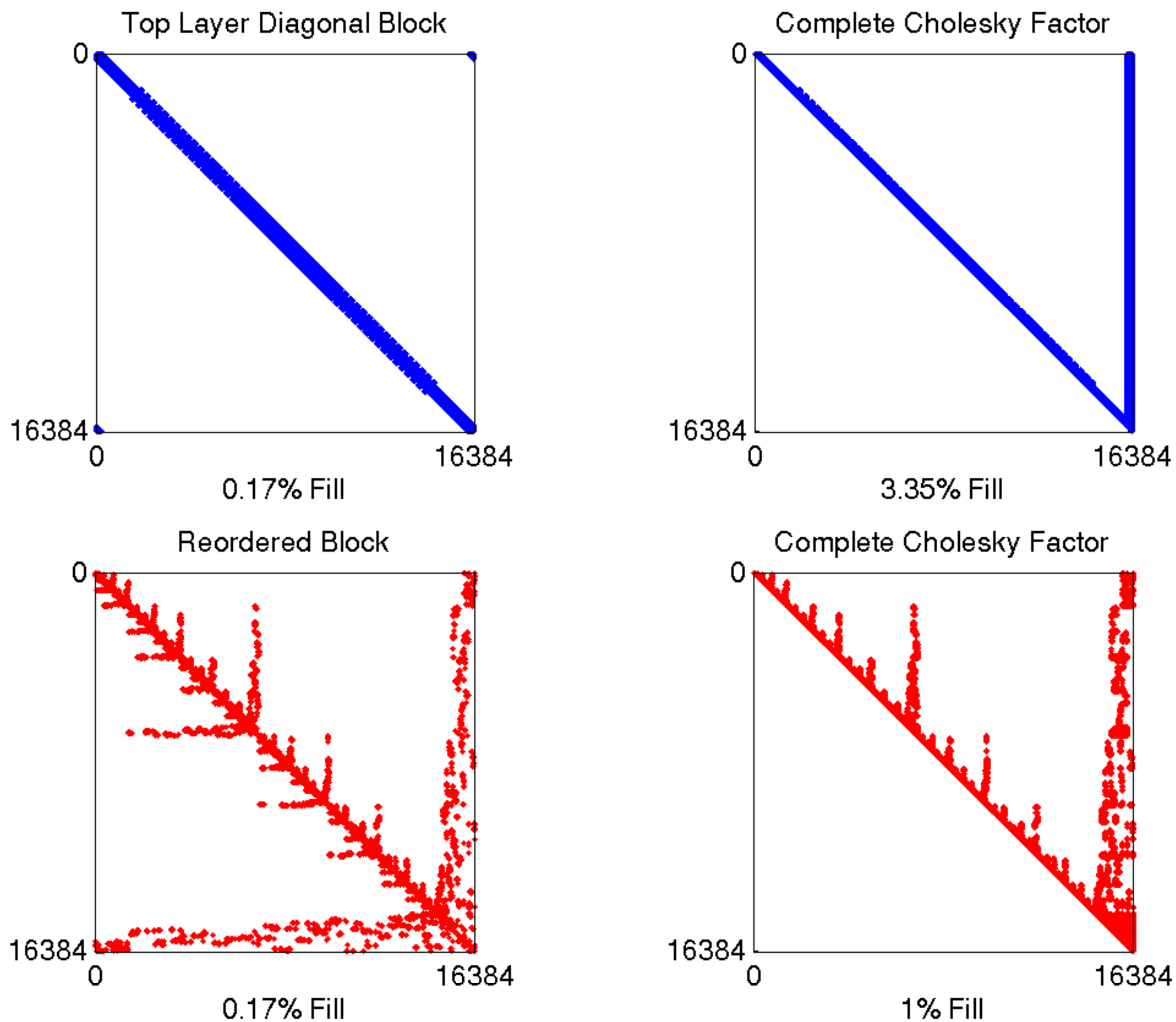


Fig.2

Fitting matrix (3 DM’s, 932 actuators/DM). D=16m.

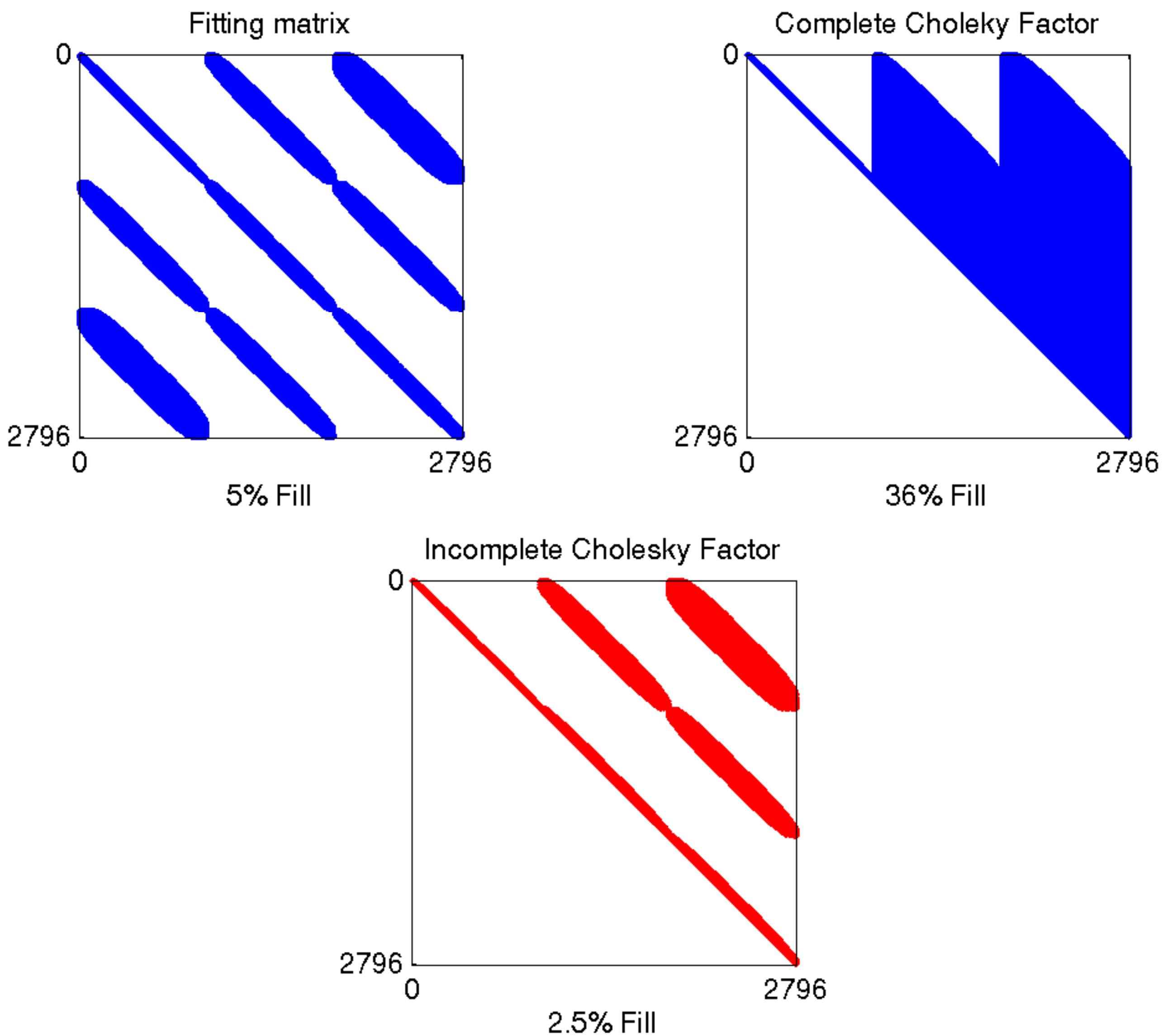


Fig.3

Estimation, 6-layer profile, 5 WFSs' using 5 NGS's. FoV diameter 100 arcsec, 1 V-cycle/CG iteration, 1 SGS iter/grid level, SNR = 20, r_0 = 25cm, $\Delta x = r_0$.

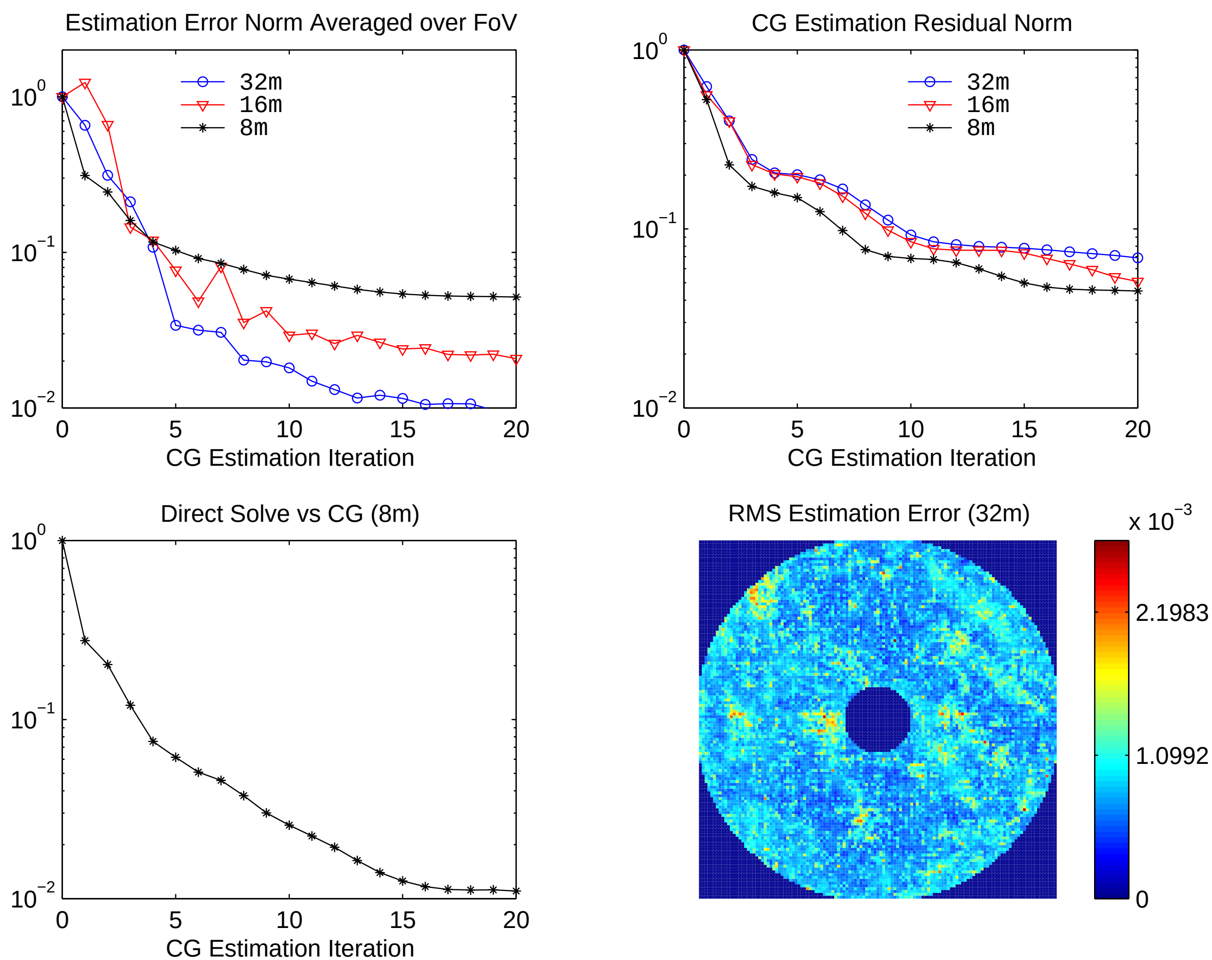


Fig.4

Fitting after 20 CG estimation iterations. Average over array of 5 $\times$ 5 observation directions. FoV diameter 100 arcsec. Incomplete Cholesky preconditioning, regularization parameter α = 1e-5.

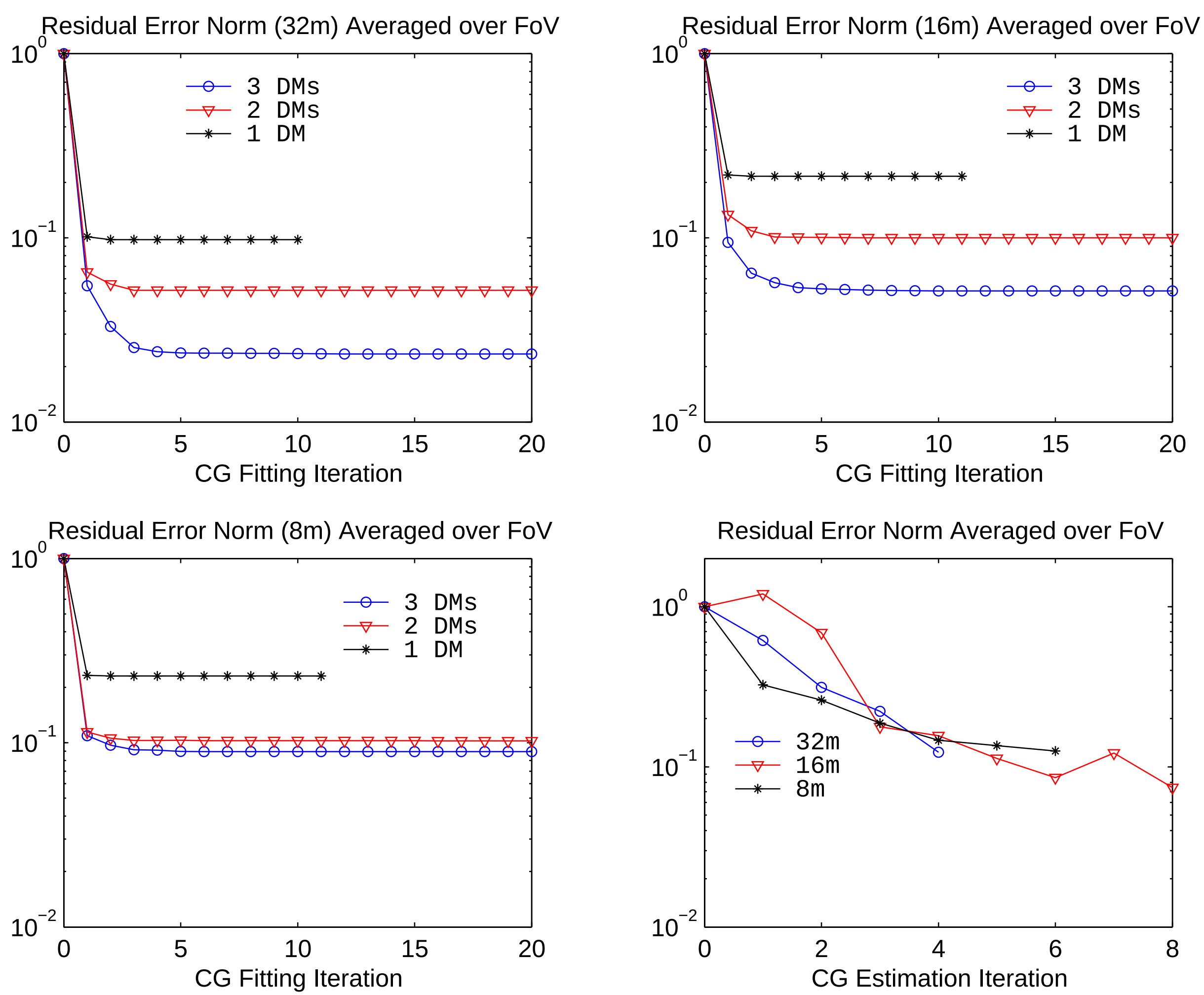


Fig.5

Fitting after 20 CG estimation iterations. Average over array of 5 $\times$ 5 observation directions. FoV diameter 100 arcsec. Incomplete Cholesky preconditioning, regularization parameter α = 1e-5.