
Keck Next Generation Adaptive Optics

Tomography Codes Comparison and Validation

Donald Gavel
UCO/Lick Observatory (gavel@ucolick.org)
April, 2007

1. Introduction and Scope

This note addresses WBS task 3.1.1.2 "Model/Tool Validation." This task involves, as stated in the System Engineering Management Plan, "Execution of a series of quantitative checks on the validity of key NGAO models and development tools, as compared to results obtained from various laboratory and sky tests with existing AO systems." In particular, we are interested in the agreement between tomography codes (3.1.1.2.1) "Understand the differences between tomography codes in use at WMKO and UCSC, modify the codes as appropriate and document the result that should be used."

2. Tomography Codes

There are three tomography codes presently in use by members of the KNGAO team:

- LAOS/TAOS written by Brent Ellerbroek, used by Chris Neyman
- A minimum variance estimation tool written by Ralf Flicker (we'll call this MVE-RF)
- TSW written by Don Gavel

The main distinction among the tomography codes is in the numerical computation algorithm that solves the minimum variance estimation problem. The solution itself should be identical under equivalent assumptions. Of course, there are a number of details in each of these complex monte-carlo codes which make it almost impossible to compare results under ideal and identical assumptions. Our approach is twofold: first independently compare the two basic implementation formulas programmed in these codes: the Moore-Penrose pseudo-inverse (LAOS/TAOS and MVE-RF) vs back projection tomography (TSW); second we directly compare simulation results from TSW to the results from the MVE-RF code reported in KAON 429[1] and comment on the differences.

3. Basic Assumptions

As has been the standard for the Keck NGAO error budgeting process, tomography error is strictly distinguished from other sources of error. Formally, the tomography error is the error in determining the plane wavefront at the ground in any field direction given noise free measurements of spherical wavefronts from a fixed constellation of guidestars. This error is most importantly affected by the turbulence strength, the spacing of guidestars, and the Cn2 profile. It is to be distinguished from generalized anisoplanatic error, which is an additional error

introduced by the discrete altitude sampling of multi-conjugate deformable mirrors. We make no assumptions about DMs but instead assume that the science wavefront in any field direction can be corrected as given by the best estimate of the line integral of turbulence volume along that direction. Tomography error is also to be distinguished from error introduced by noise in the wavefront measurements; we assume perfect, infinite signal-to-noise ratio measurements.

4. Reconstructor Algorithm Numerical Equivalence

As mentioned above, the codes use two different mathematical approaches to minimum variance estimation. These are mathematically equivalent but have differing interpretations and numerical consequences. The reader is referred to the paper by Gavel, Ammons, and Laag[2] for a full discussion. We summarize the points and conclusions here.

The codes assume that geometric optics dominates over diffractive optics in light wave propagation through weakly turbulent atmospheres. Under such conditions, there is a linear relationship between wavefront measurements and the delta-index of refraction variations in the atmospheric volume. We represent this simply by

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{x}$ are the unknown delta-index variations, $\mathbf{y}$ are the wavefront phase measurements, and $\mathbf{n}$ are the measurement noises. $\mathbf{A}$ is the linear operator (matrix) that sums up delta-index variations along rays and produces wavefront phase variations at the ground. Minimum variance estimators use a-priori second-order statistical information about the unknowns, $\left\langle \mathbf{x}\mathbf{x}^T \right\rangle = \mathbf{P}$ (i.e. a Kolmogorov turbulence spectrum) and about the measurement noises $\left\langle \mathbf{n}\mathbf{n}^T \right\rangle = \mathbf{N}$. The papers by Gavel[3,4] describe the optimization process for minimum variance and maximum likelihood estimators with and without the a-priori statistical information, and with and without noise. These papers also provide proofs of convergence and measures of convergence rates for various iterative solver algorithms.

Our objective is to minimize the wavefront error in any science direction, i.e. minimize the error in $\mathbf{y}_s = \mathbf{A}_s\mathbf{x} + \mathbf{n}$ where the subscript s denotes the science direction. It is relatively straightforward to show that a minimum variance estimate of $\mathbf{y}_s$ results from using the minimum variance estimate of $\mathbf{x}$ in this formula. There are two equivalent expressions for the minimum variance estimate of $\mathbf{x}$:

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{A}^T\left(\mathbf{A}\mathbf{P}\mathbf{A}^T + \mathbf{N}\right)^{-1}\mathbf{y} \tag{2a}$$

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A} + \mathbf{P}^{-1}\right)^{-1}\mathbf{A}^T\mathbf{N}^{-1}\mathbf{y} \tag{2b}$$

These are identical, as a consequence of the matrix-inversion lemma $\left(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A} + \mathbf{P}^{-1}\right)^{-1} = \mathbf{P} - \mathbf{P}\mathbf{A}^T\left(\mathbf{A}\mathbf{P}\mathbf{A}^T + \mathbf{N}\right)^{-1}\mathbf{A}\mathbf{P}$. Substituting this identity into (2b) will result in (2a). TSW implements formula (2a) whereas LAOS/TAOS and MVE-RF implement formula (2b). When there is no a-priori assumption about the turbulence, one sets P = identity (which results in the "least-squares" solution[3]). When there is no noise, one sets N = 0. Note that the later causes (2b) to break down,

thus any implementation of (2b) has to make some assumption about non-zero measurement noise in order for calculations to proceed.

To demonstrate that (2a) and (2b) are numerically identical we programmed both in IDL (appendix A) so that we could do one-on-one comparison on a simple case study. This demonstration code does not have the features that would make it suitable for large scale AO reconstructor simulators since it does the matrix inversions directly rather than use efficient iterative solvers. However, it is useful for illustrating the similarities and differences of formulas (2a) and (2b) in a direct manner.

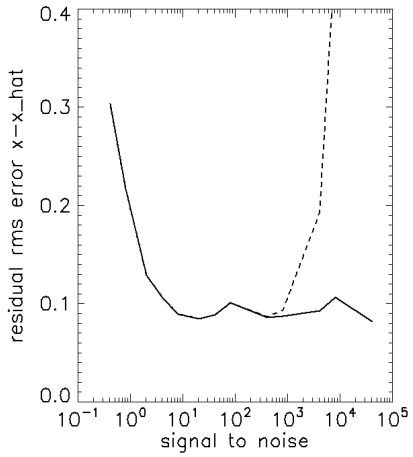


Figure 1. Residual error in the atmospheric volume estimate using method 2a (solid) and 2b (dashed). The plateau at 0.1 is due to unsensed modes in the atmosphere. The divergence of the 2b curve is due to the ill-conditioning of N^{-1} at an SNR~10^3.

We simulated a 4-layer atmosphere, an AO system with 3 guidestars arranged in a triangle on the sky, and wavefront phase sensors measuring phase on a 5x5 subaperture grids. This problem has m = 75 measurements and n = 210 unknowns, allowing for a trapezoidal shaped volume above the aperture to accommodate lines of sight to the guide stars with slope of one subaperture per layer. Figure 1 shows reconstruction error variance of the on-axis wavefront as a function of signal-to-noise ratio for the two methods. "Signal" here has been defined as the RMS phase of the planar wavefront at the on-axis science position and "Noise" is the phase-equivalent noise RMS in the wavefront sensor. Solutions are numerically identical until signal-to-noise is on the order of 1000, at which point numerical instability of (2b) results.

5. Comparision of TSW to MVE-RF

The full-up Monte-Carlo codes TSW and MVE-RF implement methods (2a) and (2b) respectively, but differ much in the details. According to R. Flicker, MVE-RF does low-order approximations to make calculations on large apertures tractable. Hence, he was able to run a very large number of asterism study cases as well as do multiple realizations of each in order to statistically average results. He used "conservative" and "ideal" cases for assumptions about the AO system configuration. As far as can be determined at this time, the "conservative" case incorporated 2 or 3 multiple conjugate DMs at layers not equivalent to those in the Cn2 model, which will introduce generalized anisoplanatism error terms. Since we want to focus on the tomography error component exclusively, we will compare TSW results only to MVE-RF

"ideal" cases. We use the model atmosphere defined in Table 1 of KAON 429. We compared results for the asterisms 4c, 5a, and 7a which are illustrated in Figure 3 of that report.

TSW implements method (2a) efficiently using a Fourier-domain preconditioned conjugate gradient (FD-PCG) iterative algorithm[3,4,5]. In our simulations, we allowed 10 iterations of PCG in each of 5 time-steps of a simple integral feedback control loop starting at $\hat{\mathbf{x}} = 0$. The turbulent layers are held fixed, i.e. we assumed no wind. In each of the cases shown, we used the same turbulence realization. It was impractical in the time scope of this study to do statistical averaging over multiple realizations, but this an option within TSW and can be done if needed.

Figure 2 shows the RMS wavefront error and Strehl results from the cases studied. This is to be compared to similar plots of MVE-RF results in Figure 10 of KAON 429. The corresponding curves clearly show the same trends and roughly the same values. From these one can make the same conclusion for example that more guide stars enable a wider field of view with a given wavefront error. 7 LGS can maintain <60nm of tomography error over a 35 arcsec radius field whereas 5 LGS can only achieve this out to about 20 arcsec. There are notable differences in these results however, which we shall point out and attempt to explain.

First of all, the TSW predicted Strehls in the wider spaced constellations are higher at science locations close to the center of the field than those of MVE-RF. This may be a consequence of the fact that TSW optimizes at each field angle independently, rather than using a compromise MCAO (multiple DM) correction over the field. I'm not sure, but this may be what is happening in MVE-RF. Alternatively, it may be the required noise regularization in formula (2b) that is limiting the MVE-RF Strehl.

Secondly, the RMS wavefront error curves of TSW are about 10-15% lower at field angles outside the nominal constellation field radii. One should not attach too much importance to this since statistical variations of the random atmosphere realization may cause this level of variation. Furthermore, and more importantly, the Strehl depends on the azimuth of the science target with respect to the constellation pattern's orientation. To give an example of this, Figure 3 shows two such azimuth lines and the corresponding TSW Strehl curves. If we evaluate at science directions on a grid over a quadrant of the science field, the resulting Strehl vs field angle plot is rather scattered (Figure 4).

Finally, the TSW RMS error curves flatten out at large field angles at about 300nm, corresponding to Strehls of about 10%. This is because we plotted the RMS error as derived from the Strehl via the Marechal approximation $S \cong \exp\left(-\sigma^2\right)$. Since Strehls only gradually decay once the diffraction-limited core is gone, the equivalent RMS error appears to flatten out. The actual RMS error over a science pupil increases rapidly once the field angle is large enough that science metapupil is no longer completely within the guide star metapupils, however this is an unfair metric since the high error variance is only concentrated in specific sections of the pupil. The well corrected area within the pupil still contributes to a, slightly broader, but diffraction-limited, PSF core; hence we plot the "equivalent" RMS wavefront error. The fact that MVE-RF curves follow analytic scaling laws out to large field angles seems to hint that MVE-RF calculations may be assuming an infinite telescope aperture.

None of these differences will change the important conclusions reached in KAON 429: more than 5 guidestars are needed to meet NGAO proposal error budget numbers in the wide field (20 and 30 arcsecond field radius) cases; and on-axis high-Strehl performance is closer to 50nm, as opposed to 29nm as stated in the proposal.

5. Comparision of TSW to LAOS/TAOS

Chris Neyman and Don Gavel ran test cases in June of last year that were intended to compare LAOS/TAOS to TSW. Although prior tests of TSW, in preparation for the 2004 SPIE meeting, had shown agreement with LAOS/TAOS, this particular test failed. LAOS/TAOS predicted a much smaller tomography error on a 60 arcsecond radius field with a 10 laser-guidestar constellation[6]. This discrepancy has since been rectified, explained by an error in assumptions during runs of TSW, and the two now agree more closely, as shown in Figure 5. The 10 LGS results are also consistent with the trends seen in KAON 429 results, where the closest similar run is constellation 9a with 38 arcsecond radius (Figure 11 in that report). Lessons learned were: a) In the earlier run, TSW's conjugate gradient solver was only run for a few (~10) iterations and hadn't converged. This small number of iterations works only in cases of small numbers of guide stars and small fields. In the latest simulations we ran for 5 time steps with 12 conjugate-gradient steps per time step and the results more indicative of a converged real time solution. b) TSW at that time was set up to compute the least-squares solution (P = identity in formula (2a)). The latest calculations are minimum variance (P = Kolmogorov turbulence) solutions.

Similar to our comparisons with MVE-RF, the TSW solutions tend to have lower RMS error near zero field angle, which is again possibly due to the fact that LAOS/TAOS may be doing some field-averaged trade-off in its fit using a finite number of multiple conjugate DMs thus mixing in some generalized anisoplanatism error, whereas TSW uses the minimum variance wavefront specific to each field direction.

Conclusions

There is reasonable agreement between the TSW code developed at UCO/Lick and the MVE-RF code used in the laser asterism study, KAON 429, in the regions of parameter space where design trades matter. There is also reasonable agreement with a test case run on LAOS/TAOS. Therefore we recommend no additional effort at this time to further improve these codes in order to achieve better agreement. It should be pointed out that TSW, at the expense of longer turn around time in calculations, seems to produce higher fidelity results than MVE-RF when finite aperture effects are important. TSW is a full-up monte carlo simulation implementing back-projection tomography (formula (2a)). It has options to oversample the wavefront to include wavefront fitting error and wavefront sensor aliasing effects as well as include or not include other AO system error effects such as generalized anisoplanatism, wavefront sensor noise, and control loop following error. These later features were explicitly "turned off" in this particular study so that only the isolated tomography error term was evaluated.

References

[1] Flicker, F., "NGAO Trade Study Report: LGS Asterism Geometry and Size," KAON 429, Nov., 2006.

http://www.oir.caltech.edu/twiki_oir/pub/Keck/NGAO/WorkProducts/KAON_429_LGSasterismTS.pdf

[2] Gavel, D., Ammons, M., Laag, E., "A Comparison of Tomography Reconstruction Techniques for MCAO and MOAO: Theory and Laboratory Experience," OSA Topical Meeting on Adaptive Optics Analysis and Methods, June, 2007.

http://lao.ucolick.org/twiki/pub/LAOLibrary/LibraryEntry78/OSApaper_2007.pdf

[3] Gavel, D., "Stability of Closed-loop Tomography for Adaptive Optics," Laboratory for Adaptive Optics Memo, Feb., 2005.

http://lao.ucolick.org/twiki/pub/LAOLibrary/LibraryEntry21/Stability_of_Closed_Loop_Tomography.pdf

[4] Gavel, D., "Stability of Closed-loop Tomography Algorithms for Adaptive Optics," OSA Topical Meeting on Adaptive Optics Analysis and Methods, June, 2005.

http://lao.ucolick.org/twiki/pub/LAOLibrary/LibraryEntry84/Stability_of_Closed_Loop_Tomography_OSA2005.pdf

[5] Gavel, D., "Tomography for MCAO Systems Using Laser Guide Stars," SPIE Astronomical Telescopes and Instrumentation, Vol. 5490, June, 2004.

http://lao.ucolick.org/twiki/pub/LAOLibrary/LibraryEntry5/TomographyGavelSPIE04.pdf

[6] Neyman, C., E-mail communication, June 5, 2006.

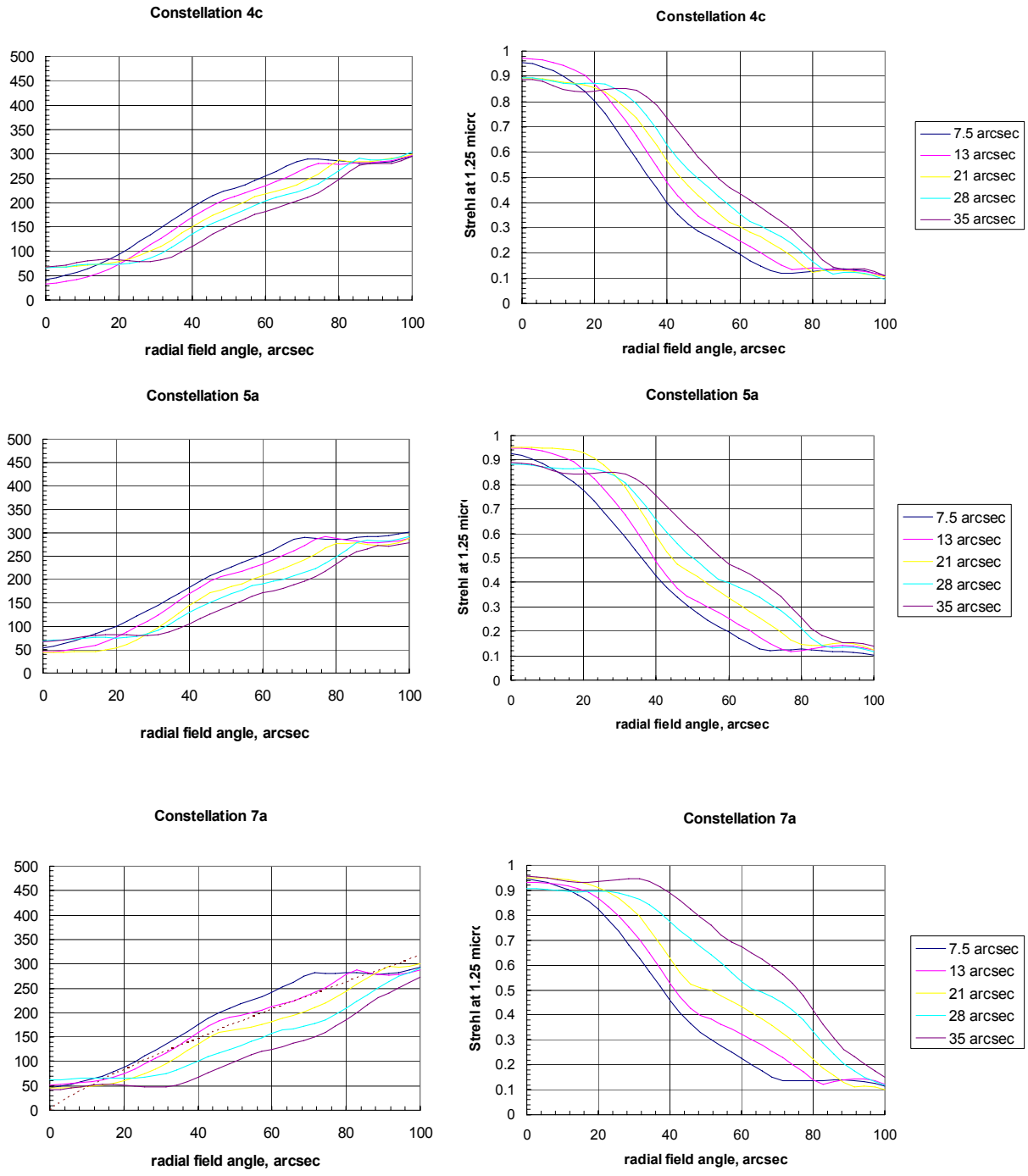


Figure 2. Tomography error computed by TSW on 4, 5, and 7 guidestar constellations. The constellation radius is given by each colored curve. The field directions are on a single radial line, with 36 equally spaced field points evaluated on each curve. In some cases, field points pass near a guide star. Compare to KAON 429 Figure 10. Charts on the left are RMS wavefront error in nm. Charts on the right are the Strehl ratio at 1.25 microns imaging wavelength.

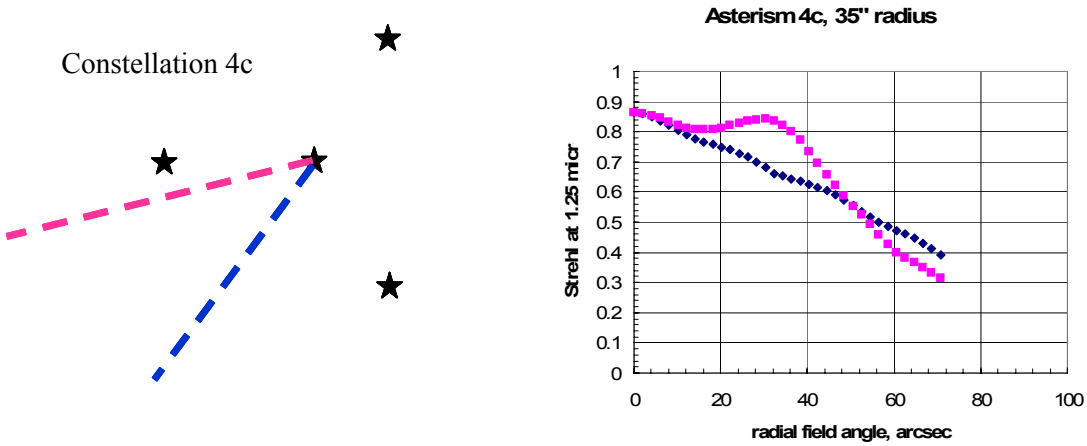


Figure 3. Tomography error differs as a function of the azimuth angle of the science target with respect to the constellation, as well as with radial position. Tomography error is lower in the vicinity of guide stars.

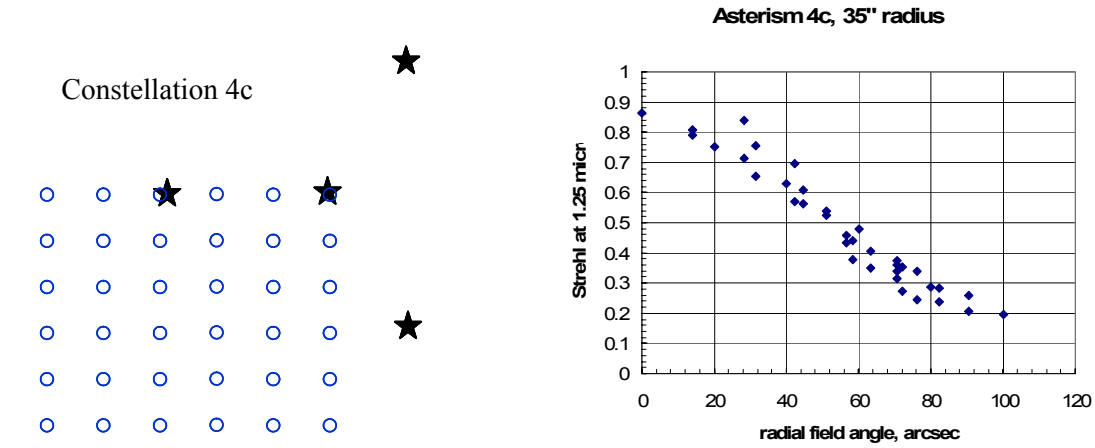
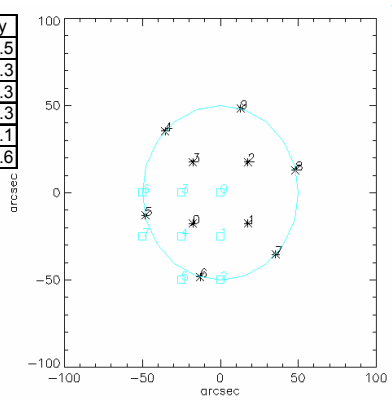


Figure 4. Scatter of tomography error over grid of science positions.

LAOS/TAOS run (6/3/06)

Only 2 armin field simulated: scaled LGS positions from image at right (CRN)
LAOS CRN 6/3/2006

Target	Radial angle, arcsec	Posn x	Posn y	Tomo only err	Tomo+fitting	fitting only
1	0	0	0	58	81	56.5
2	30	0.5	0	65	86	56.3
3	60	1	0	91	107	56.3
4	42.42640687	0.5	0.5	74	93	56.3
5	30	0	0.5	67	88	57.1
6	60	0	1	92	108	56.6



TSW run (6/3/06)

Target	Radial angle, arcsec	Posn x	Posn y	Strehl-implied nm
1	0	0	0	135.12058
2	30	0.5	0	152.57632
3	60	1	0	166.63121
4	42.42640687	0.5	0.5	178.55551
5	30	0	0.5	155.72033
6	60	0	1	203.32111

TSW runs (4/6/07)

Random seed #5

Radial_Field_ Angle/arcsec	Strehl	RMS_WFE/nm	sqrt(-ln(Strehl))/nm
0	0.971243	34.0038	33.9829
25	0.972698	33.138	33.0999
50	0.91343	62.1881	59.8647
25	0.965265	37.4411	37.4062
35.3553	0.965202	37.5384	37.4406
55.9017	0.898848	67.7631	64.967
50	0.931105	53.4512	53.1532
55.9017	0.911896	61.0391	60.4178

Random seed #13

Radial_Field_ Angle/arcsec	Strehl	RMS_WFE/nm	sqrt(-ln(Strehl))/nm
0	0.948619	45.739	45.6912
25	0.925944	55.3248	55.1835
50	0.88601	69.8863	69.2105
25	0.9475	46.2031	46.1995
35.3553	0.923829	56.0691	55.9977
55.9017	0.849733	80.8033	80.2788
50	0.863187	78.7369	76.3083
55.9017	0.839795	84.043	83.1284

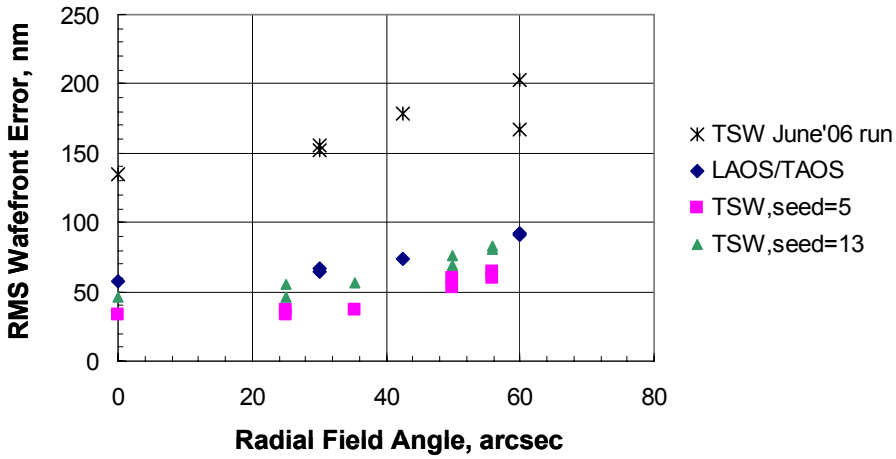


Figure 5. Comparison of LAOS/TAOS to TSW for a wide-field constellation of 10 LGS. The earlier TSW runs had not fully converged and were coded to find the least-squares solution. Steps have now been taken to assure TSW convergence to the minimum-variance solution. See text for details.

Appendix A – Code comparing method (2a) with method (2b)

This is a simple IDL code, which is independent of the detailed AO configuration assumptions made in LAOS/TAOS, MVE-RF, and TSW, and is designed to highlight the similarities and differences of equations (2a) and (2b) in tomographic reconstruction.

```
;
; OSA2007.pro
;
;  simulations to support the paper presented at the 2007 OSA
;  Topical Meeting on Analytical Methods in Astronomical Adaptive Optics
;
sigma_n_set = [10.,5.,2.,1.,.5,.2,.1,.05,.01,.005,.001,.0005,.0001]
ncases = (size(sigma_n_set))[1]
rmsAsave = fltarr(ncases)
rmsBsave = fltarr(ncases)
n_realizations = 10

for the_case = 0,ncases-1 do begin
rmsA = 0
rmsB = 0

for realization = 0,n_realizations do begin; monte-carlo realizations

ngs = 3 ; guidestars
mx = 5 ; subapertures
nl = 4 ; layers
nx = mx + 2*(nl-1) ; layer extent
n = nx*nx*nl ; number of unknowns
m = mx*mx*ngs ; number of measurements
gs_angle = [[1,0],[-1,1],[-1,-1]]
A = fltarr(nx,nx,nl,mx,mx,ngs)
for layer = 0,nl-1 do begin
  for gs = 0,ngs-1 do begin
    for ix = 0,mx-1 do begin
      for iy = 0,mx-1 do begin
        A[gs_angle[0,gs]*layer+nx/2+ix-mx/2,gs_angle[1,gs]*layer+nx/2+iy-
mx/2,layer,ix,iy,gs] = 1.
      endfor
    endfor
  endfor
endfor
yl = 2
;disp,reform(A[*,*,*,yl,yl,0]),'gs 1'
;disp,reform(A[*,*,*,yl,yl,1]),'gs 2'
;disp,reform(A[*,*,*,yl,yl,2]),'gs 3'
x = fltarr(nx,nx,nl)
y = fltarr(mx,mx,ngs)
;
; generate Kolmogorov screens
;
nn = nextpow2(nx)
r0 = 1.
du = 1.
Cn2 = [0.7,0.2,0.05,0.05]
f = screengen(nn,nn,r0,du)
acf = real(ft(f*f)*nn^2)>0 ; auto-correlation
acf = acf[nn/2-nx/2:nn/2+nx/2,nn/2-nx/2:nn/2+nx/2]
acf = zeropad(acf,nx*2,nx*2)
P = fltarr(nx,nx,nl,nx,nx,nl) ; a-priori covariance of unknowns
for layer = 0,nl-1 do begin
  for i = 0,nx-1 do begin
    for j = 0,nx-1 do begin
      for k = 0,nx-1 do begin
        for l = 0,nx-1 do begin
```

```
          P[i,j,layer,k,l,layer] = acf[i-k+nx-1,j-l+nx-1]*cn2[layer]
        endfor
      endfor
    endfor
  endfor
endfor
P = reform(P,n,n)
s = fltarr(nn,nn,nl)
A = reform(A,n,m)
for layer = 0,nl-1 do begin
  s[*,*,layer] = screengen(f,seed)*cn2[layer]
  ;s[nx/2-1,nx/2-1,layer] = layer
endfor
s = s[0:nx-1,0:nx-1,*]
for layer = 0,nl-1 do begin
  s[*,*,layer] -= average(s[*,*,layer])
endfor
; measurement noise
sigma_n = sigma_n_set[the_case]; 0.01
Nc = sigma_n*diagonal(ones(m))
noise = randomn(seed,m)*sigma_n
x = reform(s,n)
;
;  compact unused voxels
;
compact = 1
xmask = total(a,2) ge 1
if (compact) then begin
  A = A[where(xmask),*]
  x = x[where(xmask)]
  P = (P[where(xmask),*])[*,where(xmask)]
endif
;
;  forward propagation
;
y = A ## x + noise
;
wfs_y = reform(y,mx,mx,ngs)
;
;  Solutions
;   A. Filtered Back-Projection
;   B. Push Matix Inversion
;
AT = transpose(A)
Ninv = invert(Nc)
Pinv = invert(P)
R_A = P ## AT ## invert(A ## P ## AT + Nc)
R_B = invert(AT ## Ninv ## A + Pinv) ## AT ## Ninv
x_hat_A = R_A ## y
x_hat_B = R_B ## y
;
;  un-compact x_hat
;
if (compact) then begin
  u = fltarr(n)
  u[where(xmask)] = x_hat_A
  x_hat_A = u
  u = fltarr(n)
  u[where(xmask)] = x_hat_B
  x_hat_B = u
endif
;
xmask = reform(xmask,nx,nx,nl)
;
;for   layer   =   0,nl-1   do   print,'screen   rms   layer   ',strtrim(layer,2),'
',rms(s[*,*,layer])
;
normfac = total(xmask)
s_hat_A = reform(x_hat_A,nx,nx,nl)
s_err_A = s - s_hat_A
;disp,s,'s'
;disp,s_hat_A,'s_hat FBP'
```

```
;disp,s_err_A,'s_err FBP'
rmsA += sqrt(total((s_err_A*xmask)^2)/normfac)
;
s_hat_B = reform(x_hat_B,nx,nx,nl)
s_err_B = s - s_hat_B
;disp,s,'s'
;disp,s_hat_B,'s_hat PMI'
;disp,s_err_B,'s_err PMI'
rmsB += sqrt(total((s_err_B*xmask)^2)/normfac)

endfor ; realizations
rmsAsave[the_case] = rmsA/float(n_realizations)
rmsBsave[the_case] = rmsB/float(n_realizations)

print,the_case,rmsAsave[the_case],rmsBsave[the_case] & wait,.01

endfor ; cases

signal = rms(total(s,3))
!p.thick = 2
!p.charthick = 2
snr = signal/sigma_n_set
plot,snr,rmsAsave/signal,/xlog,charsize=2,linestyle   =   0,xtitle   =   'signal   to
noise',ytitle = 'residual rms error x-x_hat'
oplot,snr,rmsBsave/signal,linestyle = 2

end
```