Estimating the spatial power spectrum of residual wavefront errors from adaptive optics Monte Carlo simulations

Ralf Flicker (*rflicker@keck.hawaii.edu*)
W.M. Keck Observatory, 65-1120 Mamalahoa Hwy., Kamuela, HI 96743, USA

30 November 2006 (draft)

1 Introduction

This note addresses two possible methods for how to estimate the spatial power spectral density (PSD) function of residual turbulence errors from a numerical AO simulation. The central issue is the effect of the telescope aperture (or more generally any static aberrations) and how to separate it from the PSD of AO residual phase φ. The PSF $K(\alpha)$ and the OTF $B(\omega)$ of the system are modeled as

$$\begin{aligned} \text{PSF}: \quad K(\alpha) &= K_{\text{tel}}(\alpha) * K_\varphi(\alpha) * K_{\text{obj}}(\alpha), & (1) \\ \text{OTF}: \quad B(\omega) &= B_{\text{tel}}(\omega) \times B_\varphi(\omega) \times B_{\text{obj}}(\omega), & (2) \end{aligned}$$

where asterisk ($*$) denotes ordinary convolution. The PSF and OTF are Fourier conjugates, $K(\alpha) = \mathcal{F}[B(\omega)](\alpha)$, with $\mathcal{F}$ denoting Fourier transform, and $\{\alpha, \omega\}$ are the independent variables of angle and angular frequency in the image plane. It was assumed that the complex transmission function of the system is isoplanatic. This is an approximation, but it allows the following sequence of computations:

$$\begin{aligned} B_\varphi(\omega) &= \exp\left[-\frac{1}{2} D_\varphi(\lambda\omega)\right], & (3) \\ D_\varphi(\rho) &= 2\left[C_\varphi(0) - C_\varphi(\rho)\right], & (4) \\ \Phi_\varphi(\kappa) &= \mathcal{F}[C_\varphi(\rho)](\kappa). & (5) \end{aligned}$$

The functions and variables introduced are: the phase structure function D_φ, the phase correlation function C_φ and the phase power spectrum Φ_φ. Eqn. (3) is a consequence of the isoplanaticity property and the assumption that $\varphi(\rho)$ is a zero-mean Gaussian random variable (by virtue of the central limit theorem). Eqn. (4) follows straight from the definitions of the two functions:

$$\begin{aligned} D_\varphi(\rho) &= \langle \|\varphi(x) - \varphi(x+\rho)\|^2 \rangle, & (6) \\ C_\varphi(\rho) &= \langle \varphi(x)\varphi(x+\rho) \rangle, & (7) \end{aligned}$$

where angle brackets denote ensemble average. Finally Eqn. (5) is just a statement of the Wiener-Kinchine theorem. Inverting the first two expressions gives

$$\begin{aligned} D_\varphi(\rho) &= -2\ln\left[B_\varphi(\lambda\omega)\right], & (8) \\ C_\varphi(\rho) &= \sigma_\varphi^2 - \frac{1}{2} D_\varphi(\rho), & (9) \end{aligned}$$

where σ_φ^2 is the total residual phase variance at the given wavelength. This quantity must be provided separately by the numerical simulation. Hence, if one can obtain an estimation of the OTF B_φ, with ϕ representing the particular AO residual wavefront errors of interest, it is then straightforward to transform that into various objects characterizing the spatial statistics of the wavefront. In the following it is assumed that the object is an unresolved point source, whereby $K_{\text{obj}}(\alpha) = \delta(\alpha)$ and this part of the PSF vanishes. It is further assumed that $K_{\text{tel}}(\alpha)$ is a known quantity. Two methods for estimating B_φ are now investigated.

1.1 Method 1 (bad)

From (2) one could try to simply solve for B_φ by division:

$$B_\varphi(\omega) = \frac{B(\omega)}{B_{\mathrm{tel}}(\omega)}. \tag{10}$$

This is rather tricky for two reasons (kids, don't try this at home). First, both numerator and denominator are zero beyond the cut-off frequency of the system, and the division can only be carried out within the support of $B_{\mathrm{tel}}(\omega)$. It is reasonable to believe that this may produce a decent estimation of B_φ within a sub-domain of ω that does not reach all the way out the cut-off frequency. Proceeding with steps (8) and (9) may likewise produce an estimation of D_φ and C_φ on this sub-domain. It would not be advisable, however, to attempt the final step and compute the PSD by Eqn. (5), since there is no way to compute the Fourier transform for the valid domain without reintroducing artifacts from the boundary region close to and beyond the cut-off frequency. Some results of applying this method are shown/discussed in Section 2.

1.2 Method 2 (better)

A slightly better method might be to separate out the effects of the telescope aperture by deconvolving the PSF. This would ideally produce a K_φ that is completely free from aperture effects, and the analysis could then be taken all the way to computing the final PSD by Fourier transform. Since these PSFs can be generated without adding noise to them, a simple quadratic function minimization algorithm may be all the sophistication needed for this purpose. Switching nomenclature for a second, denote by d and $\hat{d}$ the observed data and our estimate thereof, computed isoplanatically as $\hat{d} = h * \hat{o}$, where $\hat{o}$ is the estimated object and h is the impulse response of this abstracted imaging system. Now define the quadratic cost function $\mathcal{M}$ as

$$\mathcal{M}(\hat{o}) = \|d - \hat{d}\|^2 = \sum_{i=0}^{N-1} (d - h * \hat{o})_i^2, \tag{11}$$

where the sum is over all the pixels in the image. Many function minimization algorithms rely on having an analytical expression of the gradient of the function. These include e.g. the conjugate gradient (CG), quasi-Newton and steepest descent (SD) methods. The functional derivative of $\mathcal{M}$ in the direction v is obtained as

$$\delta\mathcal{M}(\hat{o}; v) = 2(h * v)[(h * \hat{o}) - d], \tag{12}$$

where v is any admissible function. This is a general result that may be used to compute e.g. a modal derivative of $\mathcal{M}$ for any mode v. To compute a point-wise derivative we set $v = \delta$, where δ is Dirac's delta function. To ensure positivity of the result we apply the reparametrization $\hat{o} = \hat{\psi}^2$, which gives the expression

$$\delta\mathcal{M}(\hat{\psi}) = 4h \times \hat{\psi} \times [(h * \hat{\psi}^2) - d]. \tag{13}$$

This is a useful form for the algorithm, and we identify the quantities according to

$$d = K, \tag{14}$$
$$h = K_{\mathrm{tel}}, \tag{15}$$
$$\hat{\psi} = \sqrt{K_\varphi}. \tag{16}$$

To use an algorithm like e.g. a multivariate CG or quasi-Newton in IDL, all one needs to do is write a simple function that computes the cost function (11) and its derivative (13). The quasi-Newton function minimization exists as a standard IDL routine called `DFPMIN` (BFGS implementation). Unfortunately, I haven't been able to get anywhere at all using this agorithm. Maybe debugging issues, or the method simply does not work well for this particular problem. Next up, a Polak-Ribiere implementation of the conjugate gradient algorithm is provided by the `minF_conj_grad.pro` function contained in the NASA Astrolib IDL library. For reasons not yet understood, applied to the present deconvolution task, this algorithm also terminates at a zero convergence rate rather far from the minimum solution. Although to its credit, it gets there very fast.

So if the fancy stuff does not work, we fall back on the simple stuff that does. The steepest descent algorithm takes a lot of stick from the pundits for being woefully sub-optimal (seeing as this is a branch of optimization theory). But if you have a well-constrained problem and a little patience, it does the job. The algorithm is simply

$$\hat{\psi}_{k+1} = \hat{\psi}_k - g_k \delta\mathcal{M}(\hat{\psi}_k), \tag{17}$$

where g_k is a scalar feedback gain that may also be a function of the time step k, in order to speed up convergence when the gradient becomes small. What I ended up using for this study was a combination of the CG and SD algorithms, with the CG acting as a pre-processing step that gets you some of the way, leaving the remainder to the much slower SD.

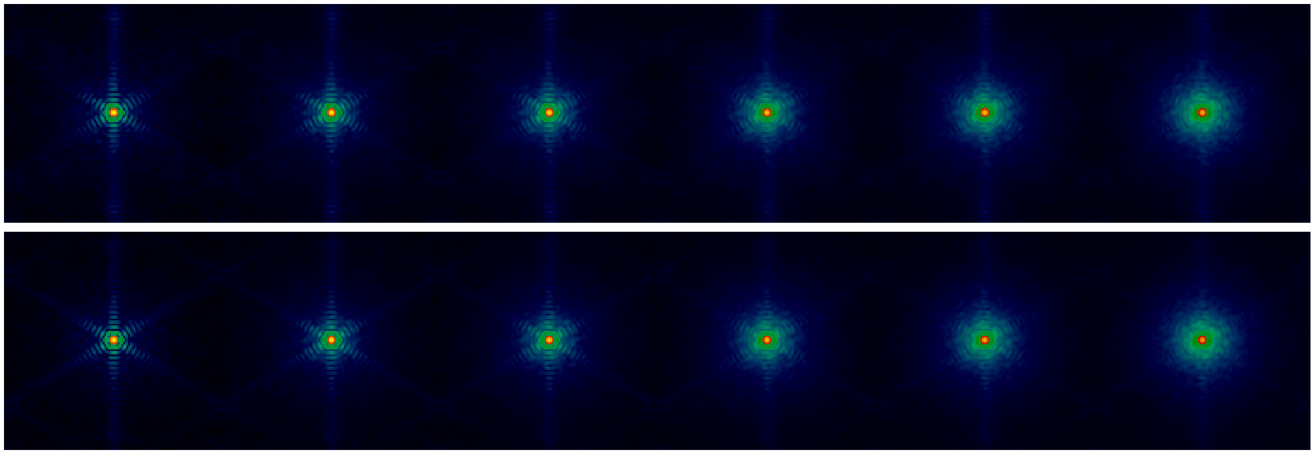

Figure 1: LTAO tomography PSFs (stretched) from the numerical simulation, for asterisms 5a (top) and 8a (bottom) and field evaluation points at distances $\{0.0, 4.2, 8.4, 5.9, 9.4, 11.9\}$ arc seconds from the central axis.

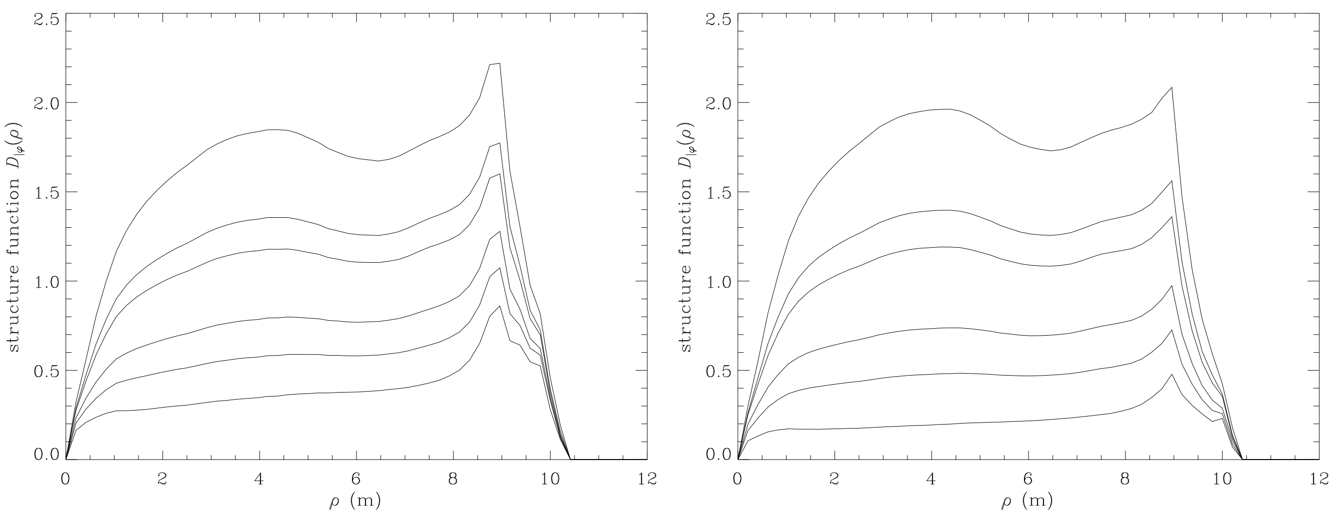

Figure 2: Radially averaged phase structure functions D_φ for asterism 5a (left) and 8a (right) computed by the simple method 1.

2 Sample numerical results

Both method 1 and 2 were applied to simulated AO PSFs produced with a Monte Carlo type simulation code described in [3]. The setup simulated a 10-m hexagonal telescope and two multi-LGS LTAO configurations with 48×48 sub-apertures. Foloowing the method of the NGAO trade study reported in Ref. [1], the code was also rigged so that the only contribution to the PSF from AO errors came from high-order tomography and anisoplanatism errors. The two asterisms compared were the 5a and 8a configurations (cf. Ref. [1] or [2]), both operating in an on-axis optimized LTAO mode. To see the influence of ordinary anisoplanatism as well as LGS tomography, five off-axis PSFs were computed in addition to the on-axis PSF, as shown in Fig. 1. The radially averaged phase structure functions obtained by deconvolving these PSFs using both methods are shown in Figs. 2 and 3. With method 1 you can clearly see the effects of the hexagonal aperture kicking around the 8-m mark, and breaking down completely after $\sim$9 m where the the support of B_{tel} runs out. With method 2 the effects of the aperture are almost entirely gone, but you can still see a trace of it around the 9-m mark, which indicates that the deconvolution was not perfect. This is also evident from looking at the phase error kernel K_φ (not shown here), which still contains some traces of the diffraction limited telescope PSF K_{tel}. The last figure 4 shows the radially averaged phase PSDs computed from the structure functions obtained with method 2.

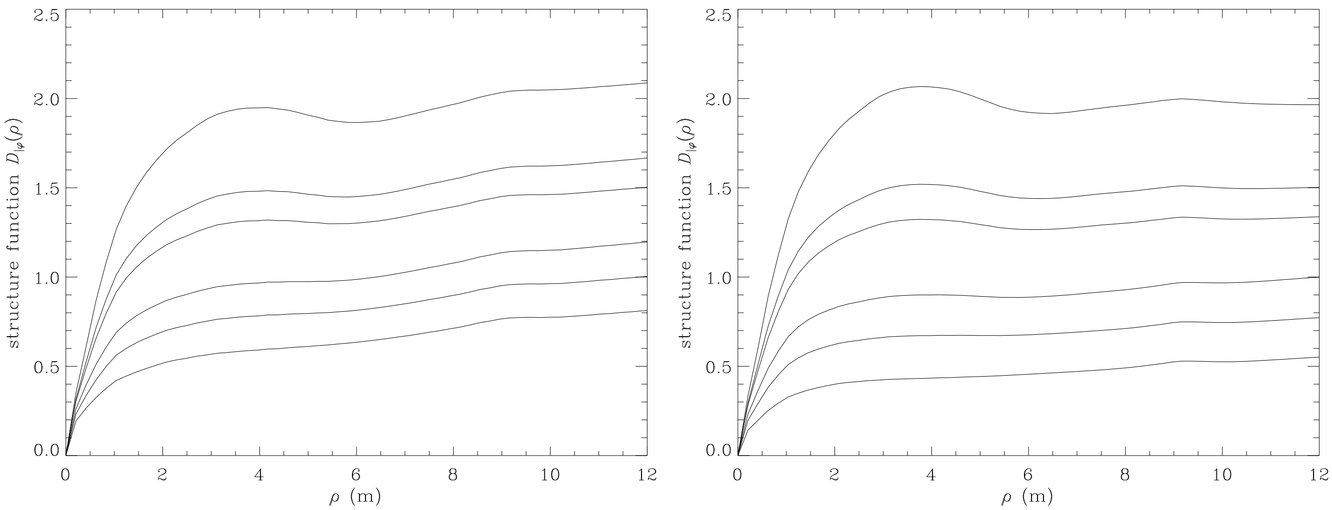

Figure 3: Radially averaged phase structure functions D_φ for asterism 5a (left) and 8a (right) computed by the Gaussian ML deconvolution method 2.

3 Notes

3.1 Estimating other AO errors

This example was drawn up for the case of separating only the telescope effects from the residual turbulence wavefront PSD. One can imagine repeating this process for separating the individual error sources of a general AO PSF, e.g. fitting error, aliasing, anisoplanatism, servo-lag, noise etc. As an example, spatial aliasing in the WFS is notoriously difficult to model analytically with good realism, while e.g. the fitting error PSD is trivial. If you arranged a simulation that contained *only* error contributions from fitting and aliasing, you could then obtain the aliasing PSD from the above approach by using the convolved telescope and fitting PSF as the known quantity in the deconvolution step.

3.2 Using with real data

To apply this method to real data rather than simulated, it may behoove us to invoke a different cost function that is modeled on the presupposed noise statistics inherent in the images. Maximum-likelihood (ML) deconvolution commonly chooses the cost function $\mathcal{M}(\hat{o}) = -\ln P(d|\hat{o})$ based on the conditional probability function. For Gaussian statistics, the form obtained in (11) is reproduced, only multiplied by $1/2\sigma^2$. For Poissonian noise, the resulting cost function is $\mathcal{M}(\hat{o}) = \sum(\hat{d} - d\ln\hat{d})$ (these are derived in e.g. Ref. [4]). Hence, the two most common statistics are trivial to insert into the algorithm discussed above and caarry out the analysis in the presence of image noise.

References

[1] R. Flicker. NGAO system design phase trade study report: LGS asterism geometry and size. Presentation given at the 2nd NGAO team meeting at Caltech, 14 Nov 2006.

[2] R. Flicker. NGAO trade study report: LGS asterism geometry and size: Kaon 4xx (WBS 3.1.2.3.3). Technical Report KAON 4xx, W.M. Keck Observatory, 14 November 2006.

[3] R. C. Flicker. Efficient first-order performance estimation for high-order adaptive optics systems. *Astron. Astrophys.*, 405:1177–1189, July 2003.

[4] R. C. Flicker and F. J. Rigaut. Anisoplanatic deconvolution of adaptive optics images. *J. Opt. Soc. Am. A*, 22:504–513, March 2005.

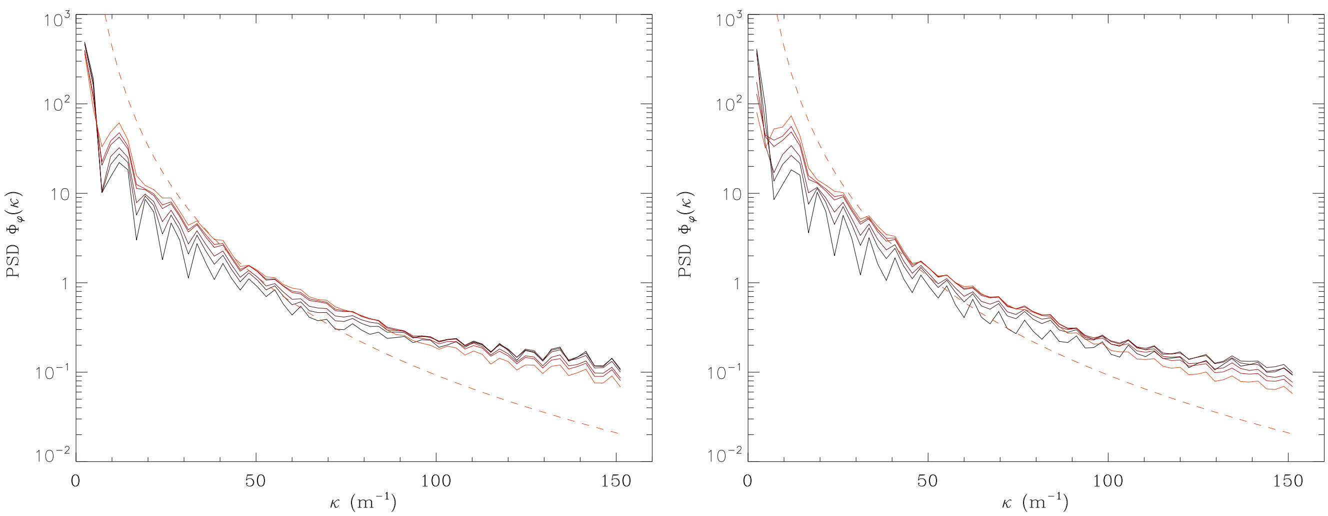

Figure 4: Radially averaged phase power spectrum Φ_φ for asterism 5a (left) and 8a (right) computed by the Gaussian ML deconvolution method 2. The dashed curve shows a $-11/3$ power law for comparison (not scaled to any physically relevant phase variance).