Working Notes on PSF reconstruction for Keck AO

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

15 April 2008

Contents

1 Introduction

This collection of working notes contains 1) summaries of the key research results from PSF reconstruction studies, and 2) the specific development undertaken in the current project for Keck AO PSF reconstruction.

1.1 Project milestones

The proposal outlined three major phases of the project, described in roughly chronological order as:

1. Initial development phase; on-axis NGS only (8 months)
    a. Review previous research on PSF reconstruction and select candidate algorithms
    b. Develop K2 AO simulation tools for the purpose of producing simulated AO telemetry data
    c. Develop TRS (telemetry recorder/server) query and analysis tools
    d. Develop prototype PSF reconstruction algorithm containing the fundamental components (fitting, aliasing, noise and bandwidth errors); test on simulated data and apply to real on-axis NGS K2 AO data
2. On-sky validation and component development phase; off-axis NGS, LGS, optical aberrations (12 months)
    a. Develop and validate angular and focal anisoplanatism components for NGS and LGS
    b. Develop static and dynamic telescope aberration components (segment figures, vibrations, instrument optical distortions); strategy for measuring them
    c. Integrated product development, preliminary deployment to routine observing
3. Final product and future development phase (4 months)
    a. Integrated product and user interface development
    b. Initial studies into PSF reconstruction techniques for future multi-beacon tomographic AO systems (i.e. MCAO, MOAO, LTAO) and extremely-high-order AO systems (applicable to, e.g., NGAO)

2 OTF calculation

2.1 The U_{ij} algorithm

First developed and implemented by Véran [28, 29] for curvature-based AO systems (PUEO), this has been the most commonly applied method also for Shack-Hartmann-based PSF reconstruction schemes to date. The method divides the phase ϕ into two orthogonal components, the vector space $\phi_\parallel$ spanned by the mirror modes (the "parallel" phase) and its orthogonal complement $\phi_\perp$ (the "orthogonal" phase), and then proceeds to compute estimates of their respective phase structure functions. The method relies on a number of assumptions:

#1. Uniform field amplitude (no scintillation)

#2. Gaussian statistics for the residual phase error

#3. Parallel and orthogonal phase components are statistically uncorrelated

#4. Structure function of the residual phase is homogeneous

#5. "Infinite bandwidth" approximation

The first three are generally valid assumptions for normal AO operation, however #4 is known to be only approximately true, and the validity might even break down in some cases (e.g. with LGS or tomography errors). The validity of the fourth assumption is something we need to investigate in more detail by simulations. Number #5 is an approximation that comes in at the modeling of the covariance matrix $\langle \epsilon\epsilon^T \rangle$ (see Sect. 3.4), and which may be improved by more detailed modeling.

2.1.1 Residual phase

With the division of spatial frequency regions into high-order ($\phi_\perp$) and low-order ($\phi_\parallel$) components, we describe the two-dimensional phase retardation of a time-dependent turbulence-aberrated wavefront by

$$\phi(\mathbf{x}, t) = \phi_\parallel(\mathbf{x}, t) + \phi_\perp(\mathbf{x}, t). \tag{1}$$

The phase correction φ introduced by the DM of the AO system can be modeled to a good approximation (see Sect. 3.2.5) as a linear sum of N_c actuator influence functions $h_i(\mathbf{x})$ according to

$$\varphi(\mathbf{x}, t) = \sum_{i=1}^{N_c} c_i(t) h_i(\mathbf{x}), \tag{2}$$

where $c_i(t)$ are modal coefficients proportional to the actuator voltages. The parallel phase $\phi_\parallel$ is by definition described in the same modal basis as the DM, so we can write

$$\phi_\parallel(\mathbf{x}, t) = \sum_{i=1}^{N_c} a_i(t) h_i(\mathbf{x}). \tag{3}$$

The PSF after AO compensation is governed by the residual phase $\epsilon(\mathbf{x}, t)$ according to

$$\begin{aligned} \epsilon(\mathbf{x}, t) &= \phi(\mathbf{x}, t) - \varphi(\mathbf{x}, t) && (4)\\ &= \phi_\perp(\mathbf{x}, t) + \sum_{i=1}^{N_c} [a_i(t) - c_i(t)] h_i(\mathbf{x}) && (5)\\ &= \phi_\perp(\mathbf{x}, t) + \epsilon_\parallel(\mathbf{x}, t), && (6) \end{aligned}$$

where we defined $\epsilon_i(t) = a_i(t) - c_i(t)$ as the modal coefficient of the residual parallel phase $\epsilon_\parallel(\mathbf{x}, t)$, and

$$\epsilon_\parallel(\mathbf{x}, t) = \sum_{i=1}^{N_c} \epsilon_i(t) h_i(\mathbf{x}). \tag{7}$$

2.1.2 Structure function

The spatial phase structure function D_ϕ is defined as

$$D_\phi(\mathbf{x}, \boldsymbol{\rho}) = \left\langle |\phi(\mathbf{x}, t) - \phi(\mathbf{x} + \boldsymbol{\rho}, t)|^2 \right\rangle, \tag{8}$$

where angular brackets denote ensemble average. Expanding the square and rearranging terms allows us to relate the structure function to the correlation function $C_\phi(\boldsymbol{\rho}) = \langle \phi(\mathbf{x}, t)\phi(\mathbf{x} + \boldsymbol{\rho}, t)\rangle$ assuming homogeneity

$$D_\phi(\boldsymbol{\rho}) = 2\left[C_\phi(0) - C_\phi(\boldsymbol{\rho})\right], \tag{9}$$

which will be eminently useful. If we assume (#3) that $\epsilon_\parallel$ and $\phi_\perp$ are statistically uncorrelated, then the structure function D_ϵ of the residual phase error is

$$D_\epsilon(\mathbf{x}, \boldsymbol{\rho}) = D_{\phi_\perp}(\mathbf{x}, \boldsymbol{\rho}) + D_{\epsilon_\parallel}(\mathbf{x}, \boldsymbol{\rho}), \tag{10}$$

where $D_{\phi_\perp}$ is the structure function of everything that the AO system does not attenuate (i.e. the fitting error), and is estimated separately (see Sect. 3.2). With the definition (7) we have that

$$D_{\epsilon_\parallel}(\mathbf{x}, \boldsymbol{\rho}) = \left\langle \left| \sum_{i=1}^{N_c} \epsilon_i(t)[h_i(\mathbf{x}) - h_i(\mathbf{x} + \boldsymbol{\rho})] \right|^2 \right\rangle \tag{11}$$

$$= \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \langle \epsilon_i(t)\epsilon_j(t)\rangle \left[h_i(\mathbf{x}) - h_i(\mathbf{x} + \boldsymbol{\rho})\right][h_j(\mathbf{x}) - h_j(\mathbf{x} + \boldsymbol{\rho})]. \tag{12}$$

One may define an aperture-averaged structure function $\bar{D}(\boldsymbol{\rho})$ according to

$$\bar{D}_\phi(\boldsymbol{\rho}) = \frac{\int d\mathbf{x}\, D_\phi(\mathbf{x}, \boldsymbol{\rho}) P(\mathbf{x}) P(\mathbf{x} + \boldsymbol{\rho})}{\int d\mathbf{x}\, P(\mathbf{x}) P(\mathbf{x} + \boldsymbol{\rho})}, \tag{13}$$

where $P(\mathbf{x})$ defines the aperture transmission function. Since the autocorrelation in the denominator is band-limited, the fraction must only be computed within the support of $\int d\mathbf{x}\, P(\mathbf{x})P(\mathbf{x} + \boldsymbol{\rho})$. The homogenized structure function $\bar{D}_{\epsilon_\parallel}$ can be written more simply as

$$\bar{D}_{\epsilon_\parallel}(\boldsymbol{\rho}) = \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \langle \epsilon_i(t)\epsilon_j(t)\rangle\, U_{ij}(\boldsymbol{\rho}), \tag{14}$$

where the U_{ij} functions are given by

$$U_{ij}(\boldsymbol{\rho}) = \frac{\int d\mathbf{x}\, \left[h_i(\mathbf{x}) - h_i(\mathbf{x} + \boldsymbol{\rho})\right]\left[h_j(\mathbf{x}) - h_j(\mathbf{x} + \boldsymbol{\rho})\right] P(\mathbf{x}) P(\mathbf{x} + \boldsymbol{\rho})}{\int d\mathbf{x}\, P(\mathbf{x}) P(\mathbf{x} + \boldsymbol{\rho})}. \tag{15}$$

The static U_{ij} functions can be computed for a given set of mirror modes $h_i(\mathbf{x})$, while the covariance matrix $\langle \epsilon_i(t)\epsilon_j(t)\rangle$ is computed from the AO telemetry stream of DM actuator commands and WFS measurements during the particular observation (see Sect. 3.4).

2.1.3 Optical transfer function

Let $K = K(\boldsymbol{\alpha}, \boldsymbol{\beta})$ be the anisoplanatic PSF as a function of image and object space angular coordinates $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, resulting from the anisotropic complex field $\psi = \psi(\mathbf{x}, \mathbf{y})$ in the image and object space spatial coordinates $(\mathbf{x}, \mathbf{y})$. For incoherent quasi-monochromatic sources it can be shown[2] that the instantaneous PSF in the far-field approximation is given by $K(\boldsymbol{\alpha}, \boldsymbol{\beta}) = |\mathcal{F}[\psi(\mathbf{x}, \mathbf{y})]|^2$, where $\mathcal{F}$ denotes the Fourier transform. Since the OTF is $B = \mathcal{F}(K)$, it can be shown by the Wiener-Khinchin theorem that the long-exposure OTF is

$$\langle B(\mathbf{u}, \mathbf{v})\rangle = \int d\mathbf{w}\, \langle \psi(\mathbf{w}, \mathbf{v})\psi^*(\mathbf{w} + \mathbf{u}, \mathbf{v})\rangle, \tag{16}$$

where $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ are angular frequency coordinates of the OTF domain, and the superscript asterisk $(^*)$ indicates complex conjugate. Assuming the field amplitude to be uniform (#1) (and hence also isoplanatic), the complex field may be described as

$$\psi(\mathbf{x}, \mathbf{y}) = P(\mathbf{x}) \exp[i\phi(\mathbf{x}, \mathbf{y})], \tag{17}$$

where P is a real-valued binary aperture transmission function and ϕ is the optical phase retardation. Combining equations (16) and (17) gives

$$\langle B(\mathbf{u}, \mathbf{v}) \rangle = \int d\mathbf{w}\, P(\mathbf{w}) P(\mathbf{w} + \mathbf{u}) \left\langle e^{i\phi(\mathbf{w}, \mathbf{v})} e^{-i\phi(\mathbf{w}+\mathbf{u}, \mathbf{v})} \right\rangle. \tag{18}$$

Assuming the phase ϕ to be a zero-mean Gaussian random variable (#2) leads to $\langle \exp(i\phi) \rangle = \exp(-\frac{1}{2}\langle \phi^2 \rangle)$, and the OTF simplifies to

$$\begin{aligned} \langle B(\mathbf{u}, \mathbf{v}) \rangle &= \int d\mathbf{w}\, P(\mathbf{w}) P(\mathbf{w} + \mathbf{u}) \exp\left[-\frac{1}{2} \langle |\phi(\mathbf{w}, \mathbf{v}) - \phi(\mathbf{w} + \mathbf{u}, \mathbf{v})|^2 \rangle \right] & (19) \\ &= \int d\mathbf{w}\, P(\mathbf{w}) P(\mathbf{w} + \mathbf{u}) \exp\left[-\frac{1}{2} D_\phi(\mathbf{w}, \mathbf{u}, \mathbf{v}) \right] & (20) \end{aligned}$$

by the definition of the structure function. If the structure function was homogeneous across the aperture plane, this wold mean that $D_\phi(\mathbf{w}, \mathbf{u}, \mathbf{v}) = D_\phi(\mathbf{u}, \mathbf{v})$, and the OTF expression simplifies further

$$\begin{aligned} \langle B(\mathbf{u}, \mathbf{v}) \rangle &= \underbrace{\exp\left[-\frac{1}{2} D_\phi(\mathbf{u}, \mathbf{v}) \right]}_{B_\phi} \underbrace{\int d\mathbf{w}\, P(\mathbf{w}) P(\mathbf{w} + \mathbf{u})}_{B_P} & (21) \\ &= B_\phi(\mathbf{u}, \mathbf{v}) \times B_P(\mathbf{u}), & (22) \end{aligned}$$

where B_ϕ is recognized as the long-exposure optical transfer function of the turbulent phase ϕ, and B_P is the static OTF of the telescope. To arrive at this result for the purpose of PSF reconstruction, Véran[28] invokes the approximation (#4) that $D_{\epsilon_\parallel} \approx \bar{D}_{\epsilon_\parallel}$, which allows the residual AO OTF to be approximated by equation (21) since $D_{\phi_\perp} = D_{\phi_\perp}(\boldsymbol{\rho})$ is explicitly homogeneous. In the case of an isoplanatic PSF or when evaluated along a single field angle, the result thus can be written

$$\begin{aligned} \langle B(\boldsymbol{\rho}/\lambda) \rangle &\approx B_{\phi_\perp}(\boldsymbol{\rho}/\lambda) \times B_{\epsilon_\parallel}(\boldsymbol{\rho}/\lambda) \times B_P(\boldsymbol{\rho}/\lambda) & (23) \\ &= \exp\left\{ -\frac{1}{2} D_{\phi_\perp}(\boldsymbol{\rho}/\lambda) \right\} \times \exp\left\{ -\frac{1}{2} \bar{D}_{\epsilon_\parallel}(\boldsymbol{\rho}/\lambda) \right\} \times (P \star P)(\boldsymbol{\rho}/\lambda), & (24) \end{aligned}$$

where $\star$ denotes convolution or autocorrelation.

2.2 The V_{ii} algorithm

A modification to the U_{ij} algorithm was recently developed by Gendron et al.[11]. A practical problem with the U_{ij} algorithm is that the computation of $\mathcal{O}(N_c^2)$ U_{ij} functions becomes very expensive for high-order AO systems. The V_{ii} method reduces the computational load to $\mathcal{O}(N_c)$ V_{ii} functions by diagonalizing the $\langle \epsilon\epsilon^T \rangle$ covariance matrix. The eigenvalue decomposition of the square symmetric matrix $\langle \epsilon\epsilon^T \rangle$ can be written

$$\langle \epsilon\epsilon^T \rangle = S \Lambda S^T, \tag{25}$$

where S is a square orthogonal matrix and $\Lambda = \mathrm{diag}(\{\lambda_i\}_{i=1}^{N_c})$ is the diagonal matrix of eigenvalues. Rewriting this as

$$\Lambda = S^T \langle \epsilon\epsilon^T \rangle S = \langle (S^T\epsilon)(S^T\epsilon)^T \rangle = \langle \eta\eta^T \rangle \tag{26}$$

results in a diagonal covariance matrix of the transformed vector $\eta = S^T\epsilon$. Conversely, substituting $\epsilon = S\eta$ into (14) and rearranging allows the structure function to be written

$$\bar{D}_{\epsilon_\parallel}(\boldsymbol{\rho}) = \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} \langle (S\eta)_i (S\eta)_j \rangle\, U_{ij}(\boldsymbol{\rho}) \tag{27}$$

$$= \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} U_{ij}(\boldsymbol{\rho}) \left(\sum_{k=1}^{N_c}\sum_{l=1}^{N_c} S_{ik}S_{jl}\langle \eta_k \eta_l \rangle \right) \tag{28}$$

$$= \sum_{k=1}^{N_c} \langle \eta_k \eta_k \rangle \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} S_{ik}S_{jk}U_{ij}(\boldsymbol{\rho}) \tag{29}$$

$$= \sum_{k=1}^{N_c} \langle \eta_k \eta_k \rangle V_k(\boldsymbol{\rho}), \tag{30}$$

where

$$V_k(\boldsymbol{\rho}) = \frac{\int d\mathbf{x} \left(\sum_{i=1}^{N_c} S_{ik} \left[h_i(\mathbf{x}) - h_i(\mathbf{x}+\boldsymbol{\rho}) \right] \right) \left(\sum_{j=1}^{N_c} S_{jk} \left[h_j(\mathbf{x}) - h_j(\mathbf{x}+\boldsymbol{\rho}) \right] \right) P(\mathbf{x}) P(\mathbf{x}+\boldsymbol{\rho})}{\int d\mathbf{x}\, P(\mathbf{x}) P(\mathbf{x}+\boldsymbol{\rho})} \tag{31}$$

$$= \frac{\int d\mathbf{x}\, \left| h'_k(\mathbf{x}) - h'_k(\mathbf{x}+\boldsymbol{\rho}) \right|^2 P(\mathbf{x}) P(\mathbf{x}+\boldsymbol{\rho})}{\int d\mathbf{x}\, P(\mathbf{x}) P(\mathbf{x}+\boldsymbol{\rho})}, \quad \text{and :} \quad h'_k(\mathbf{x}) = \sum_{i=1}^{N_c} S_{ik} h_i(\mathbf{x}), \tag{32}$$

or in vector notation $h' = S^T h$ is the transformed modal basis within which the covariance matrix $\langle \eta \eta^T \rangle$ is diagonal. This requires the computation of only N_c functions $V_k(\boldsymbol{\rho})$ (or V_{ii} in the notation of Gendron et al.). To practically compute the V_k functions we rewrite the numerator in terms of convolutions and use the Fourier transform to evaluate them:

$$V_k(\boldsymbol{\rho}) \times (P \star P) = \int d\mathbf{x}\, \left| h'_k(\mathbf{x}) - h'_k(\mathbf{x}+\boldsymbol{\rho}) \right|^2 P(\mathbf{x}) P(\mathbf{x}+\boldsymbol{\rho}) \tag{33}$$

$$= \int d\mathbf{x}\, h'^2_k(\mathbf{x}) P(\mathbf{x}) P(\mathbf{x}+\boldsymbol{\rho}) + \int d\mathbf{x}\, h'^2_k(\mathbf{x}+\boldsymbol{\rho}) P(\mathbf{x}+\boldsymbol{\rho}) P(\mathbf{x}) \tag{34}$$

$$- 2 \int d\mathbf{x}\, h'_k(\mathbf{x}) P(\mathbf{x}) h'_k(\mathbf{x}+\boldsymbol{\rho}) P(\mathbf{x}+\boldsymbol{\rho}) \tag{35}$$

$$= 2 \left[(h'^2_k P) \star P - (h'_k P) \star (h'_k P) \right] \tag{36}$$

$$= 2 \left\{ \mathcal{F}^{-1} \left[\mathcal{F}(h'^2_k P) \widetilde{P}^* - \left| \mathcal{F}(h'_k P) \right|^2 \right] \right\} \tag{37}$$

where widetilde $\widetilde{P}$ is shorthand for the Fourier transform, and the denominator is simply $P \star P = \mathcal{F}^{-1}(|\widetilde{P}|^2)$.

2.3 Non-stationary calculation

Although it might have seemed computationally implausible at one time, the diagonalization introduced in 2.2 together with Moore's Law may render the stationarity approximation unnecessary. As we shall see later on, the focal anisoplanatism structure function will in any case force us to calculate a non-stationary terms, so we might as well get some practice starting right here. Recalling the original structure function expression and applying the diagonalization of the covariance matrix, we obtain

$$D_{\epsilon_\parallel}(\mathbf{x}, \boldsymbol{\rho}) = \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \langle \epsilon_i(t) \epsilon_j(t) \rangle \left[h_i(\mathbf{x}) - h_i(\mathbf{x}+\boldsymbol{\rho}) \right] \left[h_j(\mathbf{x}) - h_j(\mathbf{x}+\boldsymbol{\rho}) \right] \tag{38}$$

$$= \sum_{k=1}^{N_c} \langle \eta_k \eta_k \rangle H'_k(\mathbf{x}, \boldsymbol{\rho}), \tag{39}$$

where

$$H'_k(\mathbf{x}, \boldsymbol{\rho}) = \left| \sum_{i=1}^{N_c} S_{ik} \left[h_i(\mathbf{x}) - h_i(\mathbf{x}+\boldsymbol{\rho}) \right] \right|^2 \tag{40}$$

$$= \left| h'_k(\mathbf{x}) - h'_k(\mathbf{x}+\boldsymbol{\rho}) \right|^2, \quad \text{and :} \quad h'_k(\mathbf{x}) = \sum_{i=1}^{N_c} S_{ik} h_i(\mathbf{x}). \tag{41}$$

And the OTF is computed explicitly as

$$\langle B(\boldsymbol{\rho}/\lambda)\rangle = \int d\mathbf{x}\, P(\mathbf{x})P(\mathbf{x}+\boldsymbol{\rho}) \exp\left\{-0.5\sum_{k=1}^{N_c}\lambda_k \left|h'_k(\mathbf{x}) - h'_k(\mathbf{x}+\boldsymbol{\rho})\right|^2\right\} \tag{42}$$

3 Component modeling

3.1 Turbulence model

Until better data comes in or until we are otherwise prompted to consider more sophisticated models, we will adopt the von Kármán turbulence model with a spatial power spectrum according to

$$\Phi(f) = \frac{0.023}{r_0^{5/3}}(f^2 + f_0^2)^{-11/6}, \tag{43}$$

where $f = |\mathbf{f}|$ and $f_0 = 1/L_0$ defines the outer scale L_0 in this model. The corresponding structure function can be written [4, 21]

$$D_\phi(\rho) = a_1 \left[a_2 - (\rho f_1)^{5/6} K_{5/6}(\rho f_1)\right] \left(\frac{L_0}{r_0}\right)^{5/3} \tag{44}$$

where $a1 \approx 0.185$, $a2 = 1.00563$ and $K_{5/6}$ is a modified Bessel function of the second kind (by some authors referred to as a modified Bessel function of the third kind, or the McDonald function). Note that here, $f_1 = 2\pi/L_0$ defines the outer scale differently than does f_0. Since the structure function can be expressed in terms of the correlation function, $D_\phi(\rho) = 2[C_\phi(0) - C_\phi(\rho)]$, we can find a useful expression for the correlation function as well:

$$C_\phi(\rho) = \frac{a_1}{2}(\rho f_1)^{5/6} K_{5/6}(\rho f_1) \left(\frac{L_0}{r_0}\right)^{5/3}. \tag{45}$$

In the instances when theoretical modeling of the atmosphere needs to be invoked on, we shall assume the Taylor hypothesis of frozen flow, and employ discrete turbulence profiles restricted to a small number of thin layers. This means that we model the structure constant as

$$C_n^2(z) = \sum_{l=1}^{N_l} c_l \delta(z - z_l), \tag{46}$$

where z_l are the layer altitudes and $c_l = f_l \mu_0$ are turbulence strengths, defined in terms of the turbulence moment $\mu_0 = 0.06\lambda^2 r_0^{-5/3}$ and the fractional weights $\sum_l f_l = 1$. Together with a wind profile v_l, the three parameters $\{f_l, z_l, v_l\}_{l=1}^{N_l}$ define a turbulence profile. Unless otherwise stated, the RD-MKR profile (see KAON #503) is consistently used.

Wiener-Khinchin theorem

Eminently useful in the U_{ij} methodology is the Wiener-Khinchin theorem, which states that the autocorrelation of a stochastic variable is equal to the Fourier transform of its power spectrum, i.e.

$$\mathcal{F}\left[\Phi(\mathbf{f})\right](\boldsymbol{\rho}) = \int d\mathbf{x} \, \langle \phi(\mathbf{x}) \, \phi(\mathbf{x} + \boldsymbol{\rho}) \rangle \;=\; C_\phi(\boldsymbol{\rho}), \tag{47}$$

provided that the Fourier transform exists. The last equality also holds when the statistics of ϕ is spatially stationary and has zero mean. While perfect in theory, there are some practical problems with applying the Wiener-Kinchin theorem to a turbulence PSD, in that the correlation function one is trying to obtain is not guaranteed to be band-limited within the computational range, i.e. the Fourier transform might not exist. We can show this to be the case for a pure Kolmogorov PSD, where simply applying the Fourier transform to Φ will roll off the correlation function at the edges of the computational grid, while in theory it should follow a 5/3-power law all the way. For a von Kármán PSD the situation improves somewhat, and gets better the smaller L_0 is, since this predicts a C_ϕ that is limited and approaches an asymptotic value at large separations, so the function naturally approximates a band-limited function eventually.

However, apart from this being a bad approximation to the theoretical C_ϕ at large separations, the total phase variance will also be underestimated as a consequence. This latter issue turns out to be the chief reason why the resulting PSF obtained in this way, by the Wiener-Khinchine theorem, will have larger Strehl ratios and smaller FWHMs than predicted by the Kolmogorov or von Kármán models. The fact that the structure function obtained by this method disagrees greatly at large separations turns out to be mostly irrelevant, since a large value of the

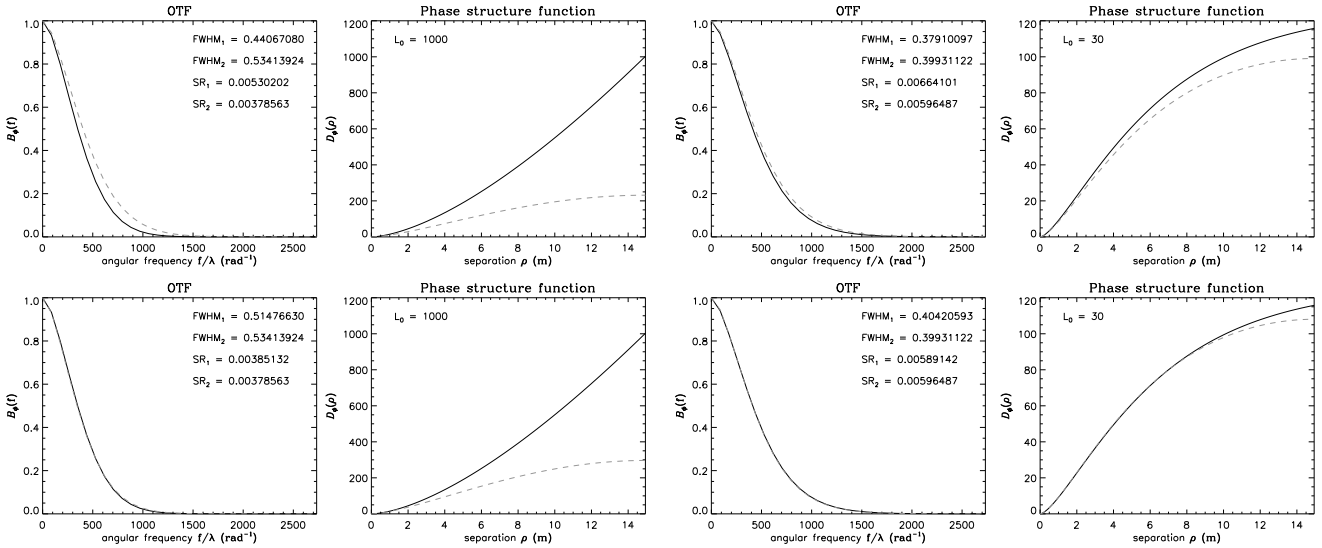

Figure 1: Turbulence model: comparing OTFs and structure functions as obtained by direc calculation of D_ϕ from equation (44) (solid black curves), and as obtained by applying the Wiener-Khinchin theorem on the PSD Φ from equation (43) (dashed grey curves). Top row: standard formulation of Φ. Bottom row: including L_0-dependent correction term.

structure function goes to a zero in the OTF very quickly by the relation (21) (already at $D_\phi = 10$, we have that $B_\phi = \exp(-5) \approx 0.0067$). So in practice we only have to be concerned with obtaining a correlation function that is a good approximation at smaller separations (which it usually is), and then we can renormalize it so that it represents the correct total phase variance in the final OTF and PSF.

I have tried to implement this renormalization here by applying the correction term $\alpha(L_0)$ to the PSD, such that $\Phi(f) \mapsto \alpha(L_0) \times \Phi(f)$ The function α was obtained numerically by a non-linear parameterized function fit to a L_0-dependent curve, which gave the following model:

$$\alpha(L_0) = a_0 - a_1 \exp(-a_2 L_0), \tag{48}$$

where the parameters are $a = [1.28925, 0.218983, 0.00326031]$. This correction term is of order unity, and is shown in Fig. 2. Examples of the structure function and the OTF with and without this correction are also shown in Fig. 1.

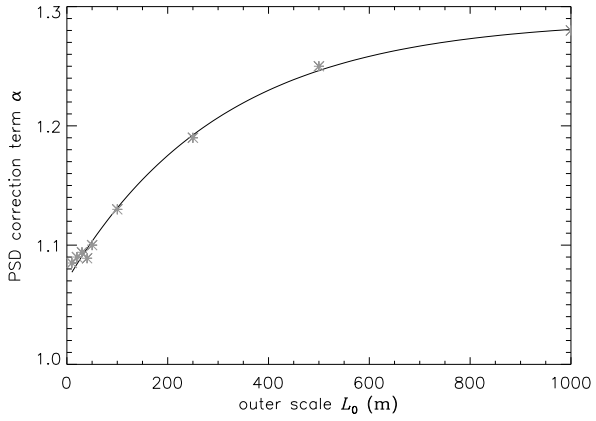

Figure 2: PSD correction term α

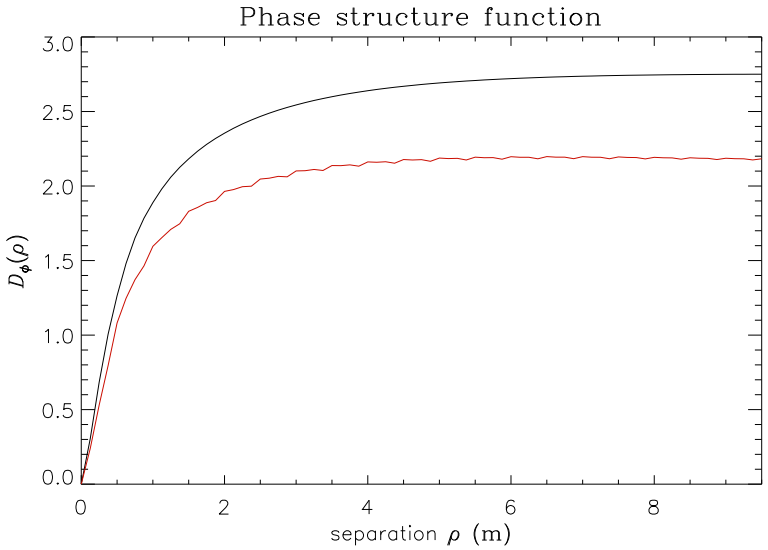

Figure 3: Structure function of the fitting error obtained with method 1. Parameter were: $D = 10$ m, $d = 0.5$ m, $r_0 = 0.5$ m, $L_0 = 30$ m. The function shown does not give the correct value for the fitting error, and getting this model to work appears to be difficult.

3.2 Fitting error

The AO fitting error comes from the non-attenuated part of the turbulence, i.e. the orthogonal phase $\phi_\perp$. Since it is the exact complement of the space spanned by the mirror modes, we can compute the fitting error by assuming perfect correction by the AO system (although there are caveats, see below). There are several approaches to computing the fitting error structure function, at varying levels of realism and computational complexity. It turns out that the most realistic models have been difficult to implement in practice, so most PSF reconstruction studies so far have opted for a simple analytical approximation that is easier to compute. For completeness, three different methods are offered here, although at the moment the first two have not yet been successfully implemented in simulation.

3.2.1 Method 1

This method models the fitting error PSD and relies on the Wiener-Khinchin theorem to obtain the structure function. Assuming that turbulence at different frequencies are statistically uncorrelated (#3), the spatial PSD Φ of the atmospheric turbulence ϕ is given by $\Phi = \Phi_\parallel + \Phi_\perp$, and hence the fitting error PSD is

$$\Phi_\perp(\mathbf{f}) = \Phi(\mathbf{f}) - \Phi_\parallel(\mathbf{f}) \tag{49}$$

$$= \Phi(\mathbf{f}) - \left\langle \left| \mathcal{F}\left[\varphi(\mathbf{x},t)\right] \right|^2 \right\rangle, \tag{50}$$

where φ is the optimal-fit DM surface that exactly cancels the parallel phase $\phi_\parallel$. This information can not be obtained from AO telemetry, but must be modeled analytically. From Sect. 2.1.1, we apply the linear DM model:

$$\varphi(\mathbf{x},t) = \sum_i a_i(t) h_i(\mathbf{x}), \tag{51}$$

where h_i are the influence functions and a_i are the ideal actuator commands. This is mathematically sound because we are observing exactly the complementary definition of the low and high spatial frequency domains defined in (1) and (2). If the influence functions are all identical in shape, which we shall assume, then $h_i(\mathbf{x}) = h(\mathbf{x} - \mathbf{x}_i)$, where $\mathbf{x}_i$ are the positions of the actuators. Defining $\mathcal{H} = |\widetilde{h}|^2$, the fitting error PSD can be written

$$\Phi_\perp(\mathbf{f}) = \Phi(\mathbf{f}) - \mathcal{H}(\mathbf{f}) \sum_{i=1}^{N_c} \sum_{j \le i} \langle a_i a_j \rangle \times \begin{cases} 2\cos[2\pi \mathbf{f}\cdot\Delta_{ij}] & , \text{ if } i \neq j \\ 1 & , \text{ if } i = j \end{cases} . \tag{52}$$

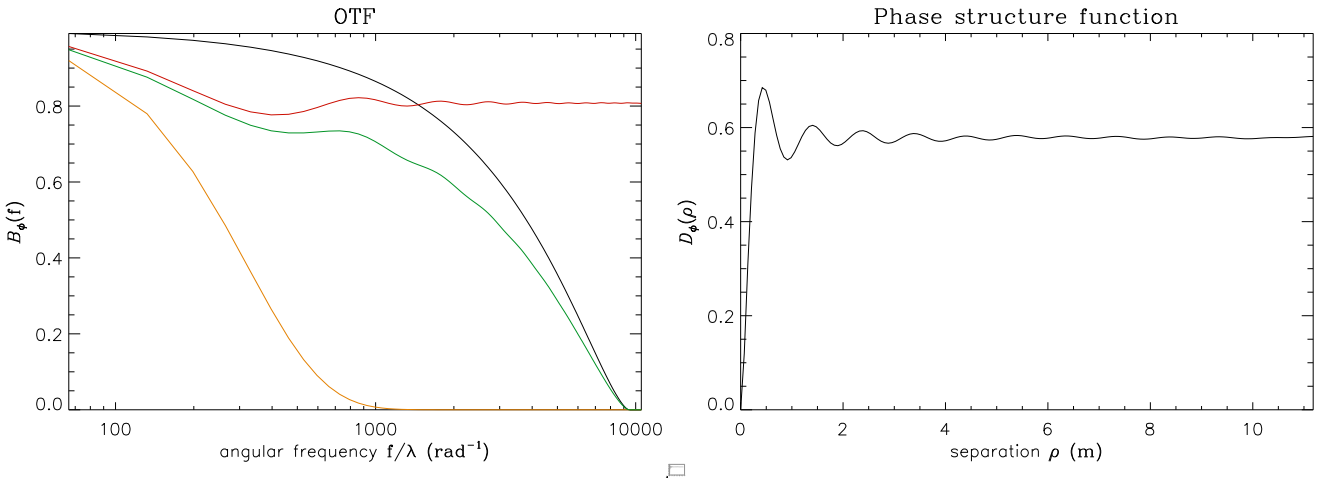

Figure 4: OTFs and structure function of the fitting error obtained with the numerically evaluated Wiener-Kinchin version of method 3. Parameter were: $D = 10$ m, $d = 0.5$ m, $r_0 = 0.6$ m, $L_0 = 30$ m, $\sigma^2_{\phi_\perp} = 0.214$ rad^2.

where $\Delta_{ij} = \mathbf{x}_i - \mathbf{x}_j$. The main difficulty here is in modeling and calculating $\langle a_i a_j \rangle$ accurately enough, since the result is formulated as a difference between to functions over a large dynamic range. If there were no cross-coupling between the actuators, one would have $\langle a_i a_j \rangle = C_\phi(\rho_{ij})$, and applying the Wiener-Khinchin theorem we obtain

$$C_{\phi_\perp}(\boldsymbol{\rho}) = [1 - \sum_{i,j} \widetilde{\mathcal{H}}(\boldsymbol{\rho} - \Delta_{i,j})] \times C_\phi(\boldsymbol{\rho}). \tag{53}$$

Fig. 3 shows two examples obtained by evaluating (53) for a bilinear spline (i.e. a pyramid) influence function (black curve) and the YAO influence function (red curve) described below with a very small coupling constant, $\beta = 0.002$. It is seen that the fitting error phase variance is greatly overestimated with respect to what is predicted for a continuous faceplate DM, which should have the constant $a_F = 0.28$ in the formula $\sigma^2_{\phi_\perp} = a_F(d/r_0)^{5/3}$. With the bilinear spline and the (decoupled) YAO influence function I obtain $a_F = 1.36$ and $a_F = 1.06$ respectively. With actuator cross-coupling one can not use the approximation $\langle a_i a_j \rangle = C_\phi(\Delta_{ij})$, and I have not obtained any results for this case yet.

3.2.2 Method 2

Another possible method is to model the structure function of the fitting error in the same way as we proposed to model the structure function of the residual AO error, using the relation $D_\phi = D_{\phi_\perp} + D_{\phi_\parallel}$. The fitting error structure function is then

$$D_{\phi_\perp}(\boldsymbol{\rho}) = D_\phi(\boldsymbol{\rho}) - \bar{D}_{\phi_\parallel}(\boldsymbol{\rho}), \tag{54}$$

where the first term is given analytically by equation (44). Projecting the parallel phase on the mirror modes as before $\phi_\parallel(\mathbf{x}) = \sum_i a_i h_i(\mathbf{x})$, we can compute the second term using the V_{ii} method (see Sect. 2.2) as

$$\bar{D}_{\phi_\parallel}(\boldsymbol{\rho}) = \sum_{i=1}^{N_c} \lambda_i V_i(\boldsymbol{\rho}) \tag{55}$$

where $\lambda_i = (S^T \langle a a^T \rangle S)_{ii}$ are the singular values of the covariance matrix $\langle a a^T \rangle$ and $V_i(\boldsymbol{\rho})$ are the equivalent of the U_{ij} function in the transformed modal basis $h' = S^T h$. While a very different computation, this method will suffer the same drawbacks as method 1 regarding the modeling of $\langle a_i a_j \rangle$. The boundaries for the complementary modal bases of $\phi_\parallel$ and $\phi_\perp$ are again observed, although now we have also packaged the pupil-averaging approximation intended for $\epsilon_\parallel$ into the analysis. I did not get this method to work yet.

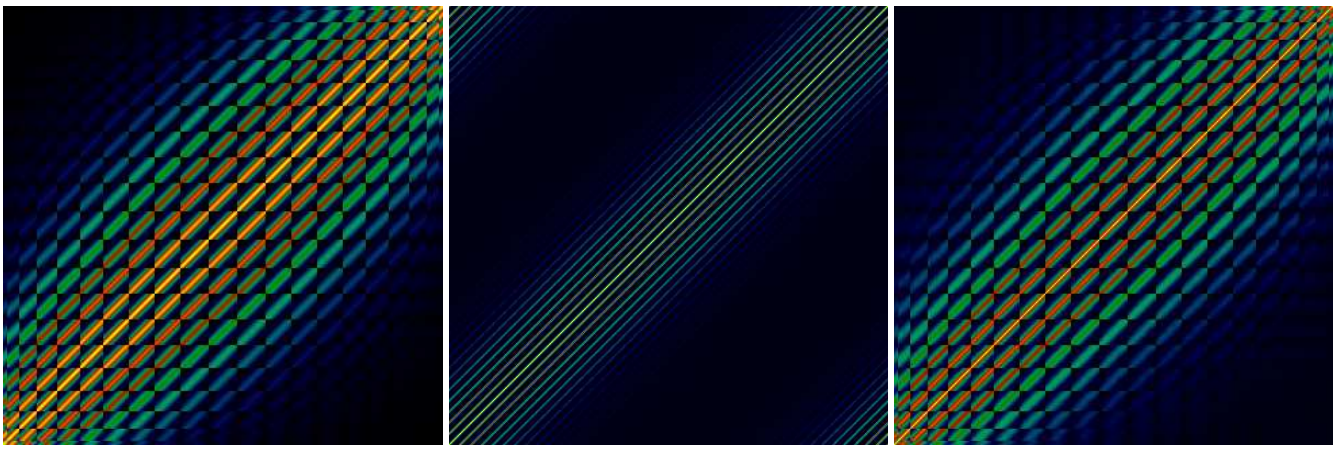

Figure 5: Covariance matrices $\langle a_i a_j \rangle$ generated with the system parameters of the Monte Carlo simulation ($L_0 = 10$ m, $r_0 = 0.5$ m). Left: theoretically computed $C_\phi(\Delta_{ij})$. Center and Right: statistically averaged $\langle a_i a_j \rangle$ from the Monte Carlo simulation covering the whole phase screen (center) and extracting only the actuators belonging to the DM.

3.2.3 Method 3

If the more realistic method 1 can not be computed accurately enough, we may have to sacrifice some mathematical rigor in order to be able to obtain a reasonably good approximation to the fitting error. One approach is to ignore the actual shape of the influence functions and the geometrical layout of actuators on the DM, and assume that the DM acts as a perfect spatial high-pass filter on the input turbulence. The fitting PSD is then uniquely defined by the spatial cut-off frequency $f_c = 1/2d$ given by the actuator separation d.

One option is to apply a binary mask $\mathcal{A}$ to the turbulence PSD and then use the Wiener-Khinchin theorem numerically to obtain the structure function. This seems to work reasonably well, although one may alternatively also compute the structure function by developing the expression analytically using relations (9) and (47). Véran derived the formula

$$D_{\phi_\perp}(\boldsymbol{\rho}) = \int_{\mathcal{A}} d\mathbf{f}\, [1 - \cos(2\pi \mathbf{f} \cdot \boldsymbol{\rho})] \Phi(\mathbf{f}) \tag{56}$$

for a square integration region where $\mathcal{A} = 1$ for $|f_x|, |f_y| > f_c$ and zero otherwise, while Jolissaint used the formula

$$D_{\phi_\perp}(\boldsymbol{\rho}) = 4\pi \int_{\mathcal{A}} d\mathbf{f}\, [1 - J_0(f\rho)] f \Phi(f) \tag{57}$$

for a circular region, $\mathcal{A} = 1$ for $f > f_c$ and 0 otherwise. Using the numerical evaluation of the PSD with a circular PSD mask and the Wiener-Khinchin theorem, I obtain a fitting error phase variance coefficient $a_F = 0.29$, while Jolissaint obtains a numerical value of 0.274, and Hardy [14] cites a coefficient of $a_F = 0.28$.

3.2.4 Monte Carlo simulation

In order to gain some understanding of the phenomena which might help us choose between different fitting error models, a Monte Carlo simulation was run where the PSDs were measured explicitly. To avoid edge-effects we use periodic random phase screens generated by the FFT method and perform the fitting of the influence functions using an extended actuator grid, covering the entire phase. Using a relatively small computational grid this still required 1600 actuators, the central 325 of which belonged to a $D = 10$ m telescope pupil, and the rest only provide boundary conditions (i.e. covering the entire screen with actuators eliminates boundaries altogether). Since the FFT screens also impose a "numerical outer scale" roughly the size of the screens $L = 20$ m, I used a small $L_0 = 10$ m for the purpose of being able to compare the covariance matrix to the theoretically calculated one $\langle a_i a_j \rangle = C_\phi(\Delta_{ij})$. The influence functions used were of the YAO type with 20% cross-coupling and $d = 0.5$ m actuator pitch in a square geometry, and $r_0 = 0.5$ m.

The simulation ran for 50000 iterations over 5000 independently generated screens, where each screen was looped over for 10 frames while shifting the screen a few pixels. The actuator covariance matrices $\langle a_i a_j \rangle$ are shown in Fig.

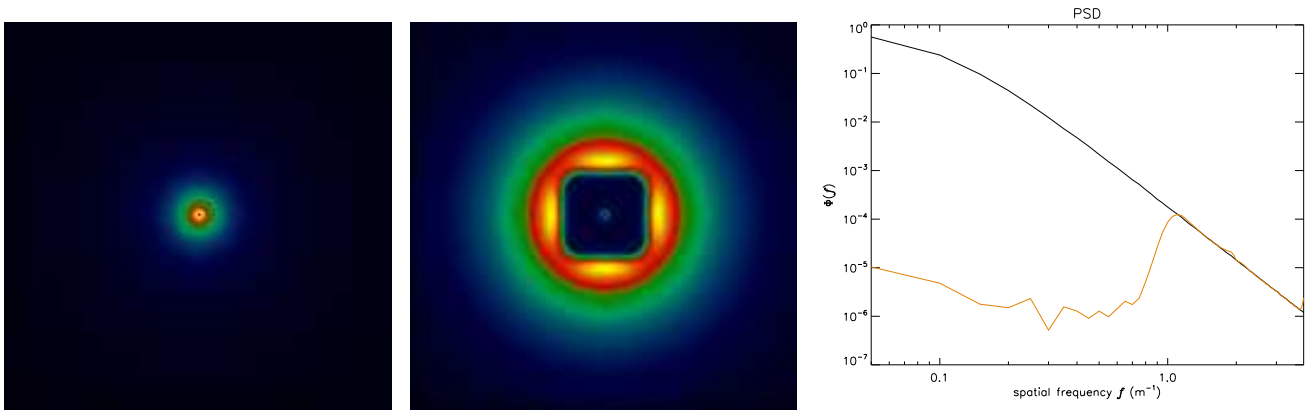

Figure 6: PSDs averaged in the Monte Carlo simulation. Left: open loop input turbulence PSD Φ_ϕ. Center: fitting error PSD $\Phi_{\phi_\perp}$. Right: cut along the x-axis comparing the pure turbulence (black curve) with the fitting error PSD (orange curve).

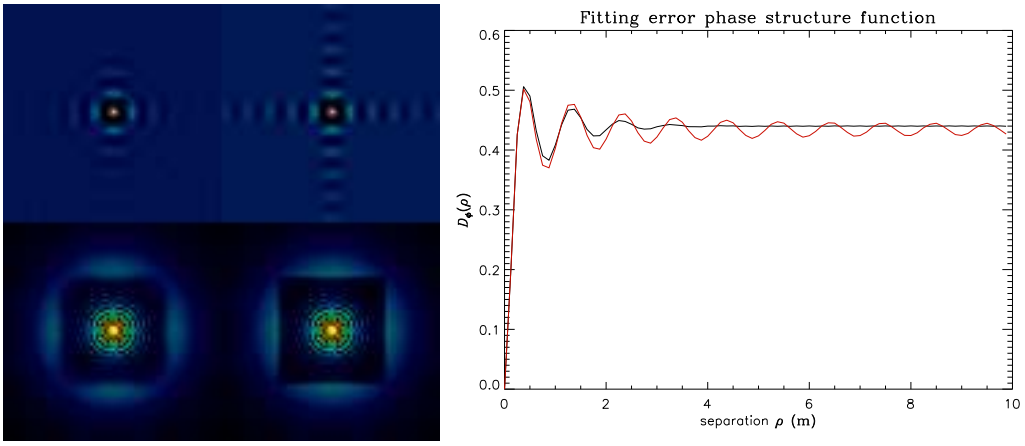

Figure 7: Fitting error OTFs, PSFs and structure functions, comparing the Monte Carlo simulation (MC) and method 3 (M3) for approximating the fitting error. Right: structure functions for the MC (black curve) and M3 (red curve) method. Top left: OTFs by the MC simulation (left) and the M3 approximation (right). Bottom left: PSFs (including the telescope kernel) from the MC sim (left) and in the M3 approximation (right).

5, with a theoretically computed $C_\phi(\Delta_{ij})$ also shown for comparison. The PSDs are shown in Fig. 6, where the total input turbulence phase variance was $\sigma^2_\phi = 10.3$ rad^2, and the residual fitting error phase variance was $\sigma^2_{\phi_\perp} = 0.224$ rad^2. Since $d/r_0 = 1$ this implies a coefficient $a_F = 0.224$ which appears to be underestimating the fitting error if the correct coefficient should be $a_F = 0.28$. That this simulation underestimates the fitting error is to be expected, since we used a relatively coarse sampling of only 4 pixels per actuator pitch (which cuts off some input power from the PSD), and there may still be some underestimation of the tilt power due to the imposed outer scale of the FFT phase screens. This discrepancy is not a major concern - one could rerun the simulation at a higher grid density and larger screen size to see if these effects go away, but it would be quite computationally expensive and time consuming so I have not done that.

Of interest to note from Fig. 6 is that the fitting error PSD does not cut off sharply at the DM cut-off frequency $f_c = 1/2d = 1$ m^{-1}, but has a somewhat gradual onset, which we should try to capture in our modeling. Fig. 7 shows fitting error OTFs, PSFs and structure functions, comparing the Monte Carlo simulation with method 3 of computing the fitting error, i.e. a binary mask on the turbulence PSD. It is seen that using a sharp mask on Φ is going to overemphasize the ringing in the structure function and the OTF, and produce a too sharp inner working region of the PSF. The effect on the Strehl of the FWHM is negligible, but the detailed structure of the PSF halo is clearly affected by the exact shape of the PSD mask, using method 3 for estimating the fitting error.

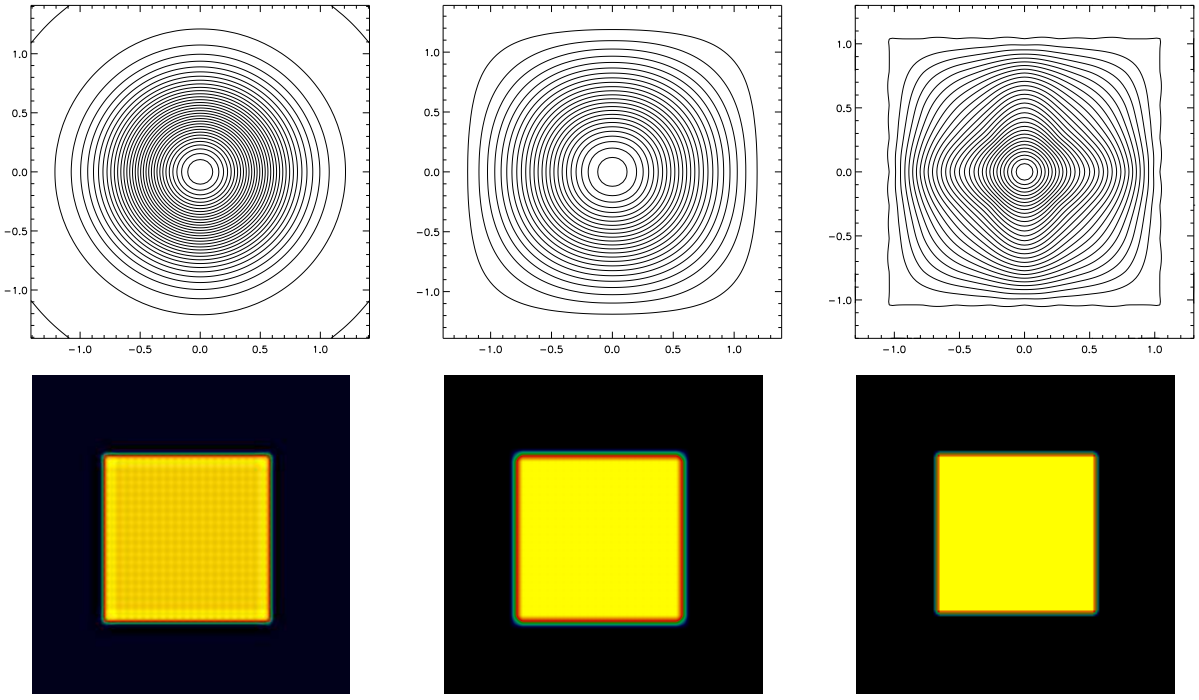

Figure 8: Influence function models, double-Gaussian Keck model (left), the parameterized YAO model (middle), and a bilinear spline model (right), showing contour plots of a single influence function (top row) and the corresponding piston mode produced by a square array of actuators (bottom row). The bilinear spline has no actuator cross-coupling, and produces a perfect piston mode; the double-Gaussian has a lot of cross-coupling and makes a pretty bad piston; the YAO model lies somewhere in between, with a modest amount of cross-coupling and a reasonably flat piston mode.

3.2.5 DM model

Different models can be applied here. M. van Dam in [27] finds that a combination of two Gaussian functions can be used for the Xinetics PZT DM in use at the Keck Observatory AO system. This influence function is given by

$$h(x,y) = 0.470\ \mu\text{m} \times \left\{ \frac{w_1}{2\pi\sigma_1^2} \exp\left[-\frac{x^2+y^2}{2\sigma_1^2}\right] + \frac{w_2}{2\pi\sigma_2^2} \exp\left[-\frac{x^2+y^2}{2\sigma_2^2}\right] \right\} \tag{58}$$

where he finds the values for the parameters to be $w_1 = 2$, $w_2 = -1$, $\sigma_1 = 0.54d$ and $\sigma_2 = 0.85d$ where d is the sub-aperture size. Advantages of this model is that it is simple to compute, and has an analytical band-limited Fourier transform that can be computed directly. Disadvantages include that it does not reproduce a piston mode very well, and it does not model the influence of neighboring actuators in square grid layout, which causes the individual influence function do deviate slightly from circular symmetry. This influence function is shown in the left panels of Fig. 8.

A different influence function model that may improve somewhat upon these two drawbacks is given by the following parameterized model which is used in the YAO simulation code:

$$h(x,y) = \begin{cases} g(x)g(y), & |x|,|y| \le p_4 \\ 0, & \text{elsewhere} \end{cases} \tag{59}$$

where

$$g(x) = 1 - |x|^{p_1} + p_3|x|^{p_2} \ln|x|, \tag{60}$$

where the parameters p_1, p_2, p_3 and p_4 are in turn parameterized functions of the coupling constant β. The coupling constant is the only free parameter of this influence function model, and is typically given a value around ~ 0.2. This model has the salient features that it produces a fairly flat and realistic piston mode, and it also progresses from a

mostly radial symmetry in the center to an x/y symmetry at the edges, which is a realistic feature that results from having a rectangular layout of actuators on the DM. The parameters of the model are given by:

$$
\begin{aligned}
a &= [0, 4.49469, 7.25509, -32.1948, 17.9493] && (61)\\
b &= [0, 2.49456, -0.65952, 8.78886, -6.23701] && (62)\\
c &= [0, 1.16136, 2.97422, -13.2381, 20.4395] && (63)\\
p_1 &= a_0 + a_1\beta + a_2\beta^2 + a_3\beta^3 && (64)\\
p_2 &= b_0 + b_1\beta + b_2\beta^2 + b_3\beta^3 && (65)\\
p_4 &= c_0 + c_1\beta + c_2\beta^2 + c_3\beta^3 && (66)\\
p_3 &= \frac{(p_4^{p_2}\beta - 1 + p_4^{-p_1})}{-\log p_4} && (67)
\end{aligned}
$$

where the a, b and c parameters were obtained (not by me) from a fitting procedure. This influence function is shown in the middle panels of Fig. 8.

As a test case I also include the simple bilinear spline influence function (rightmost panels of Fig. 8), since it has the virtues that it can produce a perfect piston mode, and can have zero cross-coupling between actuators. As a model for the DM it is not a very realistic function, but it allows us to test some analytical models (e.g. method 1) where the cross-coupling otherwise creates problems.

Recently a modified Gaussian influence function was presented (Huang et al., "Modified Gaussian influence function of deformable mirror actuators", Optics Express 16, 2008), which offers a lot of improvement over the simple or double-Gaussian model. This model has not yet been implemented or tested within the K2 AO simulation tool, and it is not clear that it would bring any significant improvement over the YAO model.

3.2.6 Action item

We still need to check the influence function model (59) against interferograms of the real DM, and tune the parameters to get the best fit. Until this can be done, we will use the YAO model with a coupling constant $\beta = 0.2$ for preliminary PSF reconstruction, which in simulations reproduces the real K2 AO interaction matrix very closely.

3.2.7 Conclusions: fitting error model

Given the evident difficulties of getting the more realistic analytical model (Method 1) to give correct results, and the oversimplified scope of Method 3, we have opted to use the numerically generated PSD mask obtained from the Monte Carlo simulations in Sect. 3.2.4 for modeling the fitting error in the PSF reconstruction algorithm. The fitting error PSD is then obtained by multiplying the input turbulence PSD by a Fourier-domain filter function $\mathcal{H}(\mathbf{f})$

$$\Phi_\perp(\mathbf{f}) = \mathcal{H}(\mathbf{f})\Phi(\mathbf{f}), \tag{68}$$

where the filter function $\mathcal{H}(\mathbf{f})$ is obtained numerically by computing both Φ and $\Phi_\perp$ explicitly in a K2 AO simulation and taking the ratio of the two. $\mathcal{H}(\mathbf{f})$ will in practice not depend on any turbulence parameters r_0, L_0 or C_n^2, but only on the shape of the influence functions h, and so we only need to pre-compute one filter function for each plate scale and PSF reconstruction grid scale that we want to use. For K2 AO this means 4-5 separate filter functions which only need to be computed once and then stored (alternatively, interpolation and scaling of a single filter function is also an option).

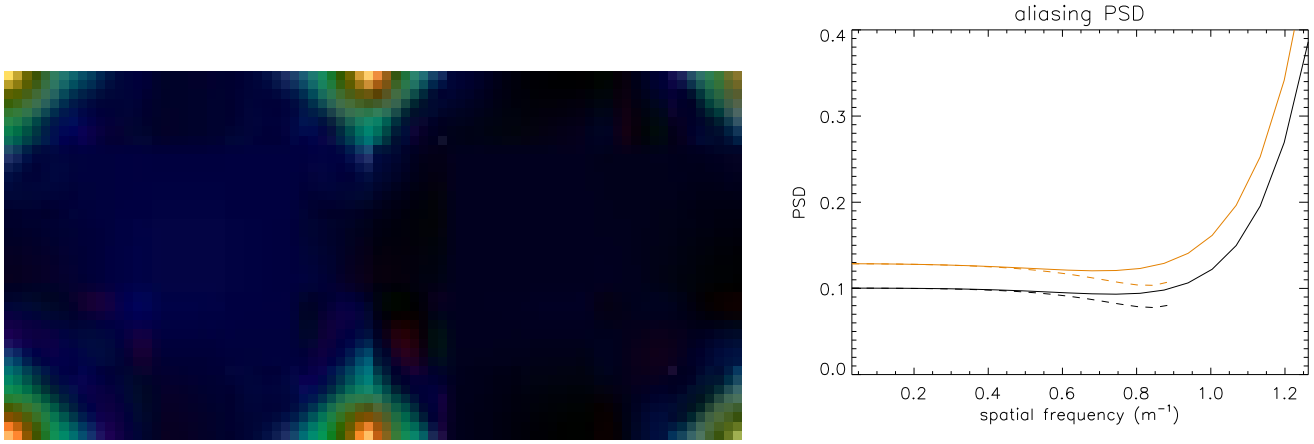

Figure 9: Aliasing PSDs by the analytical model, both for the open loop (left image; orange curves) and closed loop (image on the right; black curves) calculations.

3.3 Spatial aliasing

The expression for the WFS aliasing covariance matrix (propagated onto DM actuators) from Sect. 3.4 was:

$$\langle v_i v_j \rangle = \int d\omega \, T(\omega) E \langle r_i(\omega) r_j^*(\omega) \rangle E^T, \tag{69}$$

where, in one approximation $T(\omega) = |\mathcal{H}_n(\omega)|^2$, and in the other formulation $T(\omega) = |S(\omega)|^{-2}$. Modeling the cross-PSDs $\langle r_i(\omega) r_j^*(\omega) \rangle$ appear to be not very straightforward, and a different strategy for approximating $\langle \mathbf{v}\mathbf{v}^T \rangle$ will be explored in this section.

While modeling of aliasing in WFS domain may be tricky, there are well-developed Fourier domain models that predict the phase-error PSD resulting from WFS aliasing. The method I propose here is to use such a PSD-based aliasing model, then by the Wiener-Khinchin theorem one can obtain the corresponding phase correlation function. This correlation function can then be transformed into $\langle \mathbf{v}\mathbf{v}^T \rangle$ via the inverse of the actuator cross-coupling matrix, which can be generated numerically for a given DM actuator pattern and DM influence function. Without further ado, the formula for the spatial aliasing PSD in closed loop (assuming nothing more complicated than a leaky integrator) is given by

$$\Phi_{alias}(\mathbf{f}) = \frac{0.00575}{\mathrm{sinc}^2(\mathbf{f}d)} \times \sum_{\mathbf{m} \neq (0,0)} \frac{(\mathbf{f}^{-1} \cdot \mathbf{f}_m)^2 \, \mathrm{sinc}^2(d\mathbf{f}_m)}{(|\mathbf{f}_m|^2 + f_0^2)^{11/6}} \sum_{l=1}^{N_l} r_{0l}^{-5/3} \mathrm{sinc}^2(\mathbf{f}_m \cdot \mathbf{v}_l t_i) B_l(\mathbf{f}_m) \tag{70}$$

where $\mathbf{f}_m = (f_x - m/d, f_y = n/d)$ and d is the sub-aperture size. The function $B_l(\mathbf{f}_m)$ is the closed-loop attenuation factor, which in open loop is 1, and which for a leaky integrator has the form

$$B_l(\mathbf{f}_m) = \frac{g^2}{1 - 2a \cos b_l + a^2}, \tag{71}$$

where $a = \xi - g$, with ξ being the leak factor and g the loop gain, and $b_l = 2\pi \mathbf{f}_m \cdot \mathbf{v}_l t_i$, where t_i is the WFS integration time and $\mathbf{v}_l$ is the wind profile. This assumes the Taylor hypothesis and a layered turbulence model $\{r_{0l}, v_l\}$, as defined in Sect. 3.1. This formula is derived in Appendix 1 ("Analytical evaluations of closed-loop adaptive optics spatial power spectral densities"). It may be possible to also derive a B_l for a double-pole controller, but I have not attempted this yet.

Some examples of aliasing PSDs are shown in Fig. 9, showing both the open loop (image on the left; orange curves) and closed loop (image on the right; black curves) results for a Keck-AO-like configuration ($d = 0.56$ m). The phase variance due to aliasing has been quoted by various authors as $\sigma^2_{alias} \approx 0.3\sigma^2_{fitting}$ for a Shack-Hartmann WFS. What the closed loop analysis here shows is that this is at best roughly true, and there can be significant deviations from this rule-of-thumb in certain regions of parameter space (see Fig. 10). It is understood that the

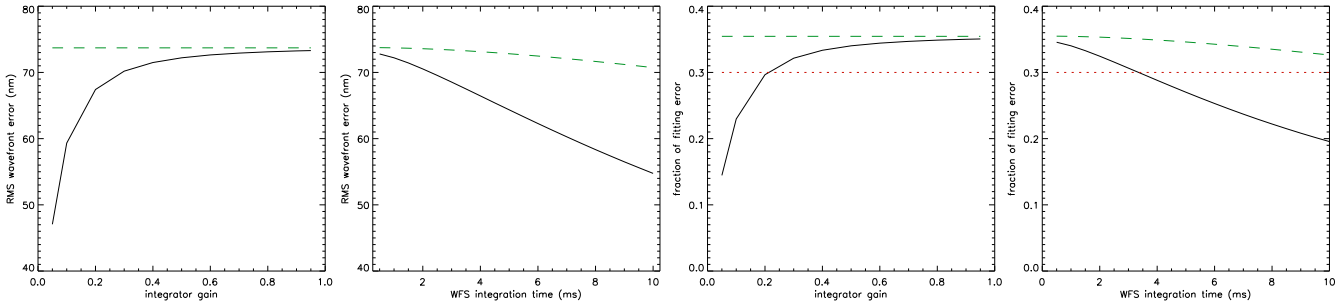

Figure 10: Aliasing wavefront error computed from analytical PSDs, as a function of WFS integration time t_i and loop gain g (solid curves are the closed-loop calculation; dashed curves are open loop).

aliasing error is, broadly speaking, anti-correlated with the servo-lag error. At lower frame rates or higher wind speeds, the orthogonal phase $\phi_\perp$ that results in aliasing will be partially averaged out while the WFS is integrating photons, and hence the aliasing error will decrease (at the expense of an increased bandwidth error). This behavior is captured realistically by the PSD model (70), as illustrated in Fig. 10. So the first virtue of using this PSD model is to obtain a more realistic estimate of the total aliasing variance to insert into the PSF reconstruction algorithm, which accounts for the closed-loop temporal filtering.

Next one can obtain the aliasing correlation function $C_{alias}(\boldsymbol{\rho})$ by Fourier transform of $\Phi_{alias}(\mathbf{f})$. Upon reformatting a little, C_{alias} becomes the phase covariance matrix defined at the actuator positions $C_{alias} = \langle \boldsymbol{\varphi}_\mathbf{v} \boldsymbol{\varphi}_\mathbf{v}^T \rangle$, where $\boldsymbol{\varphi}_\mathbf{v} = H\mathbf{v}$ (now $\mathbf{v}$ is the aliasing command vector, not the wind velocity), and H is the DM influence function matrix. Conversely, $\mathbf{v} = H^+ \boldsymbol{\varphi}_\mathbf{v}$, and we finally obtain the aliasing covariance matrix as

$$\langle \mathbf{v}\mathbf{v}^T \rangle = H^+ \langle \boldsymbol{\varphi}_\mathbf{v} \boldsymbol{\varphi}_\mathbf{v}^T \rangle (H^+)^T. \tag{72}$$

The virtue of doing it this way, rather than just taking the aliasing PSD Φ_{alias} and computing the OTF directly by the sequence established in Sect. 2.1.3, is that error now gets propagated onto the mirror modes properly. In a direct calculation of the OTF from the PSD we would not be observing the definition of the parallel phase $\varphi_\parallel$, as the modal basis of the DM modes. All the same, this $\langle \mathbf{v}\mathbf{v}^T \rangle$ is a model-based approximation to the real aliasing contribution, since we basically tossed out the originally derived expression in (69) and put in place a whole new estimation of the term.

3.3.1 Conclusions: aliasing model

Some verification of this modeling approach has been undertaken. Monte Carlo simulations of the original WFS aliasing covariance matrix $\langle \mathbf{r}\mathbf{r}^T \rangle$ reveal a diagonally dominant matrix with a very short correlation length. Propagated onto actuators by the reconstruction matrix E we obtain something very similar to $\langle \mathbf{v}\mathbf{v}^T \rangle$ as obtained with the PSD-based method. The model 72 will be used tentativaly. The impact of the cross-term with the residual error $\langle \boldsymbol{\epsilon}\mathbf{v}^T \rangle$, as well as the significance of temporal filtering, is investigated in Sect. 3.4.

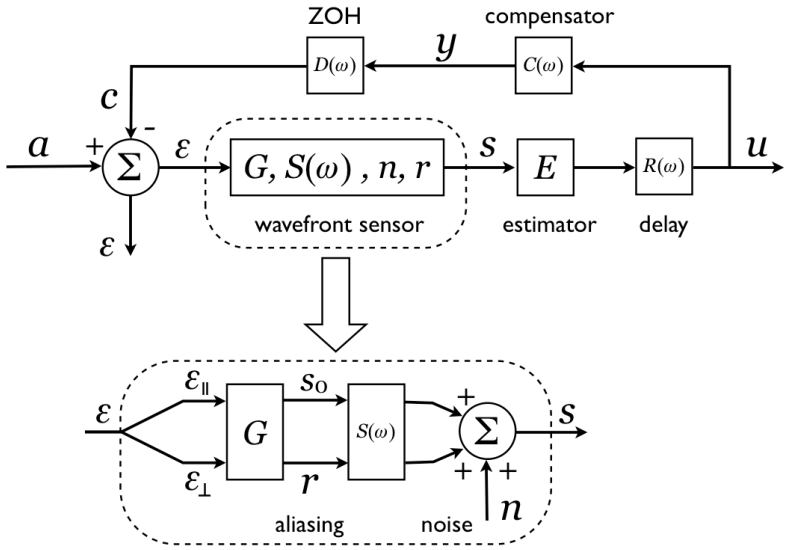

Figure 11: Schematic of AO temporal control loop (top) and a blowup of the WFS model (bottom). E and G are spatial matrix mappings, independent of time, while S, R, C and D are temporal filters, generally matrix-valued but assumed in this analysis to be scalar filters

3.4 Covariance matrix $\langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle$

Exact temporal modeling of the residual mode covariance matrix $\langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle$ turns out to be problematic, and most implementations so far have applied varying levels of approximation to this aspect of the system modeling. Referring to the generic AO block diagram in Fig. 11, we can examine the loop dynamics in terms of the parallel phase modal coefficients $\mathbf{a}$ (input turbulence), $\mathbf{c}$ (DM mode) and $\boldsymbol{\epsilon}$ (residual error). The other variables are: $\mathbf{s}_0$ is the (ideal) noise- and aliasing-free WFS measurement; $\mathbf{n}$ is the noise and $\mathbf{r}$ is the aliasing, and $\mathbf{s}$ is the actual WFS measurement; the $\mathbf{u}$ and $\mathbf{y}$ variables are used in defining the compensator (see below), and they are also variables which we have access to from AO telemetry data.

A subtlety to be aware of is that the low- and high-spatial-frequency modal sets $\phi_\parallel$ and $\phi_\perp$ were defined by the DM modal basis $h_i(\mathbf{x})$, and strictly speaking this need not be equivalent to the subsets of ϕ that are aliasing-free and produce aliasing in the WFS, respectively (if the influence functions h_i were sinc functions, the domains would be the same). Hence, the aliasing $\mathbf{r}$ can not be said to have come solely from $\phi_\perp$, since the DM influence functions themselves may produce a (very small) amount of aliasing. In practice, however, in a Fried-geometry Shack-Hartmann AO system, the modal basis of the DM influence functions and the basis that is aliasing-free are very close to each other, and to a first-order approximation we can say that aliasing only arises from $\phi_\perp$ (and hence $\epsilon_\perp$).

To repeat the basic result from Sect. 2.1.2, we are trying to obtain the structure function

$$\bar{D}_{\epsilon_\parallel}(\boldsymbol{\rho}) = \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \langle \epsilon_i(t)\epsilon_j(t)\rangle \, U_{ij}(\boldsymbol{\rho}). \tag{73}$$

Since we can not access directly the vector $\boldsymbol{\epsilon}$, we need to express the covariance matrix $\langle \epsilon\epsilon^T \rangle$ in other variables which can be obtained from AO telemetry data, such as $\mathbf{u}$, $\mathbf{y}$ and $\mathbf{s}$. Analyzing the block diagrams in Fig. 11 of the closed-loop AO system and the WFS model, we can write down the following relationships:

$$\widetilde{\boldsymbol{\epsilon}} = \widetilde{\mathbf{a}} - \widetilde{\mathbf{c}} \tag{74}$$

$$\widetilde{\mathbf{c}} = D(\omega)C(\omega)\widetilde{\mathbf{u}} \tag{75}$$

$$\widetilde{\mathbf{u}} = R(\omega)E\widetilde{\mathbf{s}} \tag{76}$$

$$\widetilde{\mathbf{s}} = S(\omega)(\widetilde{\mathbf{s}}_0 + \widetilde{\mathbf{r}}) + \widetilde{\mathbf{n}} \tag{77}$$

$$\widetilde{\mathbf{s}}_0 = G\widetilde{\boldsymbol{\epsilon}}_\parallel \tag{78}$$

where all the vectors are represented in Fourier domain as functions of the temporal frequency ω, and we omitted the frequency argument for simplicity. Definitions of the temporal filters are given in e.g. [12] (we will specify them later for the Keck AO system). Putting it all together and rearranging terms we obtain the general closed-loop relation:

$$[I + DCRESG]\widetilde{\boldsymbol{\epsilon}} = \widetilde{\mathbf{a}} - DCRE\widetilde{\mathbf{n}} - DCRES\widetilde{\mathbf{r}}, \tag{79}$$

where G is the interaction matrix and E is the wavefront estimator. In the general case, the expression $I+DCRESG$ is matrix valued, and a matrix inverse must be computed for every value of ω in order to solve for $\boldsymbol{\epsilon}$. This appears to be impractical, and a common approach to approximate the solution is to make the following additional assumptions: 1) the transfer functions D, C, R and S are all scalar-valued, and 2) $EG \approx I$. Condition #1 can be met if there is no mix-and-match of different WFS types and DM controllers that have different sets of parameters (no modal control). Condition #2 must generally be treated as an approximation, in particular for the standard MAP or SVD estimator this requirement is known to be violated. But accepting it as an approximation, we can separate the temporal from the spatial operators in (79) and solve for $\boldsymbol{\epsilon}$ simply by scalar division. The result is:

$$\widetilde{\boldsymbol{\epsilon}}(\omega) = \mathcal{H}_\epsilon(\omega)\widetilde{\mathbf{a}}(\omega) - \mathcal{H}_n(\omega)\widetilde{\mathbf{m}}(\omega) - \mathcal{H}_\epsilon(\omega)\widetilde{\mathbf{v}}(\omega), \tag{80}$$

where $\mathcal{H}_\epsilon = 1/(1+\mathcal{H})$ is the error transfer function, $\mathcal{H}_n = DCR/(1+\mathcal{H})$ is the noise transfer function and $\mathcal{H} = DCRS$ is the open loop transfer function. We also introduced the actuator command vectors $\mathbf{m} = E\mathbf{n}$ and $\mathbf{v} = E\mathbf{r}$ for the reconstructed noise and aliasing signals. When integrated over the Real line, Parseval's theorem allows us to equate $\langle \epsilon_i \epsilon_j \rangle = \langle \widetilde{\epsilon}_i \widetilde{\epsilon}_j^* \rangle$, which we can invoke here since we know the PSDs must be band-limited, and there is no power in the zero-frequency component (we remove the DC term). The covariance matrix elements are then

$$\begin{aligned} \langle \epsilon_i \epsilon_j \rangle &= \int d\omega \, \langle \widetilde{\epsilon}_i(\omega) \widetilde{\epsilon}_j^*(\omega) \rangle & (81) \\ &= \int d\omega \, |\mathcal{H}_\epsilon|^2 \, \langle \widetilde{a}_i(\omega) \widetilde{a}_j^*(\omega) \rangle + \int d\omega \, |\mathcal{H}_n|^2 \, \langle \widetilde{m}_i(\omega) \widetilde{m}_j^*(\omega) \rangle + \int d\omega \, |\mathcal{H}_\epsilon|^2 \, \langle \widetilde{v}_i(\omega) \widetilde{v}_j^*(\omega) \rangle, & (82) \end{aligned}$$

where it was assumed that $\mathbf{a}$, $\mathbf{m}$ and $\mathbf{v}$ are all statistically uncorrelated (which they are). This is a nice expression for theoretical modeling, but not very useful for PSF reconstruction since it contains the open loop turbulence vector $\mathbf{a}$ which is unknown. The covariance matrix needs to be reformulated in terms of $\mathbf{u}$.

Alternatively we can explore the direct relation between $\boldsymbol{\epsilon}$ and $\mathbf{u}$ that is given by equations (76)-(78), which reads

$$\widetilde{\mathbf{u}} = RE\left[S(G\widetilde{\boldsymbol{\epsilon}} + \widetilde{\mathbf{r}}) + \widetilde{\mathbf{n}}\right] \tag{83}$$

Without (much) loss of generality we may again assume that the temporal filters are scalar-valued, which gives

$$EG\widetilde{\boldsymbol{\epsilon}} = \frac{1}{RS}\widetilde{\mathbf{u}} - \frac{1}{S}\widetilde{\mathbf{m}} - \widetilde{\mathbf{v}}. \tag{84}$$

This expression can be used to evaluate the covariance matrix $\langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle$ at varying levels of approximation, as explored below.

Method 1

At the most drastic level of approximation we ignore the temporal filtering, and also invoke as before the approximation $EG \approx I$, which gives the simplest possible form:

$$\boldsymbol{\epsilon} = \mathbf{u} - \mathbf{m} - \mathbf{v}. \tag{85}$$

In evaluating the covariance matrix one will obtain cross-terms, and it may be a matter of taste or prudence how to arrange the expression before computing the covariance, which determines whether you will have cross-terms with $\boldsymbol{\epsilon}$ or with $\mathbf{u}$. The traditional choice has been to compute the cross-terms in terms of $\boldsymbol{\epsilon}$, and following this pattern we obtain

$$\langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle = \langle \mathbf{u}\mathbf{u}^T \rangle - \langle \mathbf{m}\mathbf{m}^T \rangle - \langle \mathbf{v}\mathbf{v}^T \rangle - 2\langle \boldsymbol{\epsilon}\mathbf{m}^T \rangle - 2\langle \boldsymbol{\epsilon}\mathbf{v}^T \rangle. \tag{86}$$

Véran [28] argues that the aliasing terms can be approximated as $-\langle \mathbf{v}\mathbf{v}^T \rangle - 2\langle \boldsymbol{\epsilon}\mathbf{v}^T \rangle \approx \langle \mathbf{v}\mathbf{v}^T \rangle$, and assuming the noise to be uncorrelated with ϵ gives the solution:

$$\langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle = \langle \mathbf{u}\mathbf{u}^T \rangle - \langle \mathbf{m}\mathbf{m}^T \rangle + \langle \mathbf{v}\mathbf{v}^T \rangle. \tag{87}$$

The validity of the reducing the aliasing terms into one according to the above formula is looked into in Sect. 3.3.

Method 2

Keeping the modeling of the aliasing cross-terms as in Method 1, but going back to include the previously omitted temporal filters, gives the somewhat more complicated expression

$$\langle \epsilon_i \epsilon_j \rangle = \langle v_i v_j \rangle - \langle m_i m_j \rangle \int d\omega \, \mathrm{sinc}^{-2}(\omega T) + \int d\omega \, \langle \widetilde{u}_i(\omega) \widetilde{u}_j^*(\omega) \rangle \, \mathrm{sinc}^{-2}(\omega T), \tag{88}$$

where it was assumed that R is a pure delay τ, i.e. $R(\omega) = \exp(-2\pi i \tau)$ and $S(\omega) = \mathrm{sinc}(\omega T) \exp(-i\pi\omega T)$, so that taking the modulus squared eliminates the complex exponentials. (see Sect. 3.3 for an evaluation of the cross-term). It was also assumed that the temporal noise PSD is constant (i.e. white noise – see Sect. 3.6), and the covariance term could be was extracted from the integral. While this expression invokes different (fewer) approximations, it is instead rather more demanding to compute since we now need cross-PSDs for all the time series of $\mathbf{u}$. This might not be so onerous as it sounds though. With 349 actuators we then "only" have a total of 61075 PSDs $\langle \widetilde{u}_i(\omega) \widetilde{u}_j^*(\omega) \rangle$ to compute. This is probably quite manageable, since they are computed by 1-dimensional Fourier transforms, they can be computed off-line, and we probably don't need very high resolution either (to be verified), since we don't expect a lot of power at the lowest temporal frequencies. A real-time calculation it is not, however.

Method 3

This is variation that can be applied equivalently to either Method #1 or Method #2, since it involves only a different treatment of the spatial operators. The analysis will derived here for Method #1. Up until now we have not cared about the actual form of the reconstructor E, but we made the statement that $EG \approx I$ as invoked in both Method #1 and #2 might be a pretty bad approximation. Assuming a specific form for E we can develop expression (84) further. Currently in use at the Keck AO system is a MAP estimator of the form

$$E_{\mathrm{MAP}} = (G^T C_n^{-1} G + C_c^{-1})^{-1} G^T C_n^{-1}, \tag{89}$$

where $C_n = \langle \mathbf{n}\mathbf{n}^T \rangle$ is the noise covariance matrix and $C_c = H^+ \langle \boldsymbol{\phi}_\parallel \boldsymbol{\phi}_\parallel^T \rangle (H^+)^T$ is the open loop phase covariance matrix mapped onto actuators. Instead of invoking $EG \approx I$ in (84) we expand the expression, defining for brevity the matrix $A = G^T C_n^{-1} G$

$$A\boldsymbol{\epsilon} = (A + C_c^{-1})\mathbf{u} - G^T C_n^{-1}(\mathbf{n} + \mathbf{r}). \tag{90}$$

Assuming that the inverse A^{-1} exists (A is usually singular, and the inverse can be approximated by SVD), we have the new expression for the residual mode covariance matrix:

$$\boldsymbol{\epsilon} = (I + A^{-1} C_c^{-1})\mathbf{u} - E_{\mathrm{GM}}(\mathbf{n} + \mathbf{r}), \tag{91}$$

where $E_{\mathrm{GM}} = A^{-1} G^T C_n^{-1} = (G^T C_n^{-1} G)^{-1} G^T C_n^{-1}$ is the Gauss-Markov estimator. This is only subtly different from (84) in Method #2, as the approximation now takes place in a different location: instead of invoking $EG \approx I$ we have to approximate A^{-1} by filtering its singular modes in the inversion. I have no idea if this is a better or a poorer representation, but it would be a simple thing to compare it to Method #1, and all those cross-PSDs are going to be computed anyway we can also do the comparison with Method #2.

Control law

Van Dam [27] describes the DM control law currently in use by the recurrence equations, with u and y as defined in Fig. 11:

$$e(n) = -we(n-1) + k_{\mathrm{DM}} u(n), \tag{92}$$

$$y(n) = ly(n-1) + e(n), \tag{93}$$

where $w = 0.25$ and the leak factor l is close to 1. The transfer function is given in z-domain by

$$\mathcal{H}_{\mathrm{DM}}(z) = \frac{k_{\mathrm{DM}} z^2}{(z-l)(z+w)}, \tag{94}$$

and the corresponding Fourier domain transfer function is given by $C(\omega) = \mathcal{H}_{\mathrm{DM}}(z = e^{-2\pi i \omega T})$.

3.5 Tilt structure function

The tip/tilt (TT) structure function is stationary and can be taken outside of the OTF integral, even in the non-stationary computation of the rest of the OTF. The structure function is given by the quadratic form

$$D_{tt}(\boldsymbol{\rho}) = \frac{4}{R^2}\left[\Gamma_{11}^2\rho_x^2 + \Gamma_{22}^2\rho_y^2 + 2\Gamma_{12}\rho_y\rho_x\right], \tag{95}$$

where the gamma coefficients are elements of the covariance matrix

$$\Gamma = \begin{bmatrix} \langle c_2^2\rangle & \langle c_2 c_3\rangle \\ \langle c_3 c_2\rangle & \langle c_3^2\rangle \end{bmatrix}, \tag{96}$$

and $c_{1,2}$ are the estimated residual tip/tilt coefficients in the phase error expansion $\varphi_{tt}(\mathbf{x},t) = c_2(t)Z_2(\mathbf{x}) + c_3(t)Z_3(\mathbf{x})$, where the Zernike modes are defined as $Z_2(\mathbf{x}) = 2x$ and $Z_3(\mathbf{x}) = 2y$. Since D_{tt} is independent of $\mathbf{x}$ it can be taken outside of the OTF integral, under the assumption that ϵ and $c_{2,3}$ are orthogonal, i.e. TT has been perfectly filtered from the DM commands. This may not be completely true, and an alternative approach could include TT in the non-stationary calculation by appending TT to the DM set $h(\mathbf{x})$ and ϵ. In the separable case, the quadratic form of D_{tt} makes the TT OTF into a Gaussian function, with the width and orientation specified by the covariance matrix Γ:

$$\langle B_{tt}(\boldsymbol{\rho}/\lambda)\rangle = \exp\left\{-\frac{1}{2}D_{tt}(\boldsymbol{\rho})\right\} \tag{97}$$

3.6 Noise covariance matrix

There are approximately as many different noise estimation techniques as there are PSF reconstruction projects, but the principal techniques commonly employed are:

1. Modeling from first principles, detector physics and photon statistics (e.g. Véran et al. 1997 [28], Jolissaint et al. 2004 [15])
2. Curl calculations (e.g. Tyler 2000 [24] and tOSC reports no. TR-816 and TR-881; also Hardy 1998 [14])
3. Temporal autocorrelation of control loop variables (e.g. Gendron & Léna 1995 [13], Egner 2004 [??])

The first two methods both suffer from the problematic of accounting for sub-apertures that deviate from the ideal, for instance by being partially illuminated (at Keck this is a function of time, since the WFS rotates with respect to the telescope pupil). For the physical modeling you have to be able to know and model rather accurately these imperfections, which is probably not feasible to do, specially if they are not constant. The curl methods can only estimate an average noise level, and again one must be careful to exclude sub-apertures that may be corrupted by external effects that produce a non-zero curl not related to either the noise or the aliasing. The curl method is probably not useful for us in this application.

The average sub-aperture intensity levels are reported by TRS, so could be used as input in method 1. But there is the additional problem of estimating the spot size with undersampled sub-apertures, since a given pixel-noise level will translate into varying centroid noise levels depending on the spot size and the centroid gain setting. This complicates calculations from first principles considerably by introducing another poorly known quantity. There are proposed methods for estimating the spot size (e.g. TT dithering or slope discrepancy [25]), but their accuracy and viability are somewhat disputed, and may work even worse with LGS.

The autocorrelation method also has the virtue that it characterizes each sub-aperture from the telemetry data itself without any a priori assumptions, so that particular effects causing the noise level to deviate from the norm are taken into consideration. It becomes less straightforward in closed loop where the turbulence correlation is much smaller, and to make the method work a somewhat ad hoc term is applied (Hamilton 1994). It should be considered whether the method could instead be applied directly to the actuator commands rather than sub-aperture slopes. While this requires a lot more computation (the matrix is no longer diagonal), we would get the propagated noise directly and avoid the problem of needing to estimate spot sizes, and also avoid the closed-loop complication.

Physical modeling

Will not do it this way.

Curl calculation

(we will not do it this way, but I wrote this down anyway while I was thinking about it)
Vector algebra shows that $\nabla \times \nabla\phi = 0$ (curl grad $\phi = 0$) in a simply connected domain ϕ, which we can exploit as a constraint on the slope measurement $\mathbf{s} = \nabla\phi$. In matrix notation, the curl $\mathbf{q}$ can be computed on a slope vector $\mathbf{s}$ by a linear operator Q

$$\mathbf{q} = Q\mathbf{s}, \tag{98}$$

where the curl matrix Q has the dimensions $N_r \times N_m$, the number of curls N_r times the number of measurements N_m. When implemented in practice, one should also exclude sub-apertures with partial illumination or otherwise affected to produce a non-zero curl that does not originate from the aliased wavefront measurement or detector noise. *(Question: does the segmented primary mirror or the segment aberrations prevent a curl-based noise estimation from being carried out?)* The WFS slope vector can be written $\mathbf{s} = \mathbf{s}_0 + \mathbf{r} + \mathbf{n}$, where $\mathbf{s}_0$ is the aliasing-free turbulence measurement, $\mathbf{r}$ is the aliasing vector, and $\mathbf{n}$ is the noise. Since $Q\mathbf{s}_0 = 0$ and $\langle r_i n_j \rangle = 0 \; \forall \; \{i, j\}$ we have that

$$\mathbf{q} = Q\mathbf{r} + Q\mathbf{n} \tag{99}$$

and

$$\sigma_q^2 = \langle \mathbf{q}^T\mathbf{q} \rangle = \langle \mathbf{r}^T Q^T Q \mathbf{r} \rangle + \langle \mathbf{n}^T Q^T Q \mathbf{n} \rangle \tag{100}$$

$$= \langle \mathbf{r}^T Q^T Q \mathbf{r} \rangle + \mathrm{Tr}(Q^T Q)\sigma_n^2, \tag{101}$$

where σ_n^2 is the noise variance in a single sub-aperture, and we assumed that the noise is uncorrelated between sub-apertures, i.e. the covariance matrix $\langle \mathbf{n}\mathbf{n}^T \rangle = \sigma_n^2 I$ is proportional to the identity matrix. The aliasing may be correlated across the WFS however. There may be different ways of using this information. If we had a model for the aliasing curl variance $\sigma_{q_r}^2 = \langle \mathbf{r}^T Q^T Q \mathbf{r} \rangle$, we could measure the total curl variance σ_q^2 and solve for the noise variance:

$$\sigma_n^2 = \frac{\sigma_q^2 - \sigma_{q_r}^2}{\mathrm{Tr}(Q^T Q)}. \tag{102}$$

Autocorrelation

We investigate whether this can be done on the actuators, to get the propagated noise directly. The actuator cross-correlation as a function of time l is

$$\begin{aligned} V_{xy}(l) &= N^{-1} \sum_{i=l}^{N} (x_i - \langle x \rangle)(y_{i-l} - \langle y \rangle) && (103) \\ &= \langle (x_i - \langle x \rangle)(y_{i-l} - \langle y \rangle) \rangle \,. && (104) \end{aligned}$$

If the loop frame rate is much faster than the turbulence decorrelation time, then the noise level can be approximated by

$$\begin{aligned} V_{xy}(0) - V_{xy}(1) &= \langle (x_0 - \langle x \rangle)(y_0 - \langle y \rangle) \rangle - \langle (x_1 - \langle x \rangle)(y_0 - \langle y \rangle) \rangle && (105) \\ &= \langle (x_0 - x_1)(y_0 - \langle y \rangle) \rangle && (106) \\ &= \langle (x_0 - x_1) y_0 \rangle \,, && (107) \end{aligned}$$

where it was assumed that $\langle x_0 \rangle = \langle x_1 \rangle$ by ergodicity. The result 107 is straightforward to compute from TRS data by shifting the telemetry stream one time step, taking the difference and matrix multiplying with itself to give the actuator noise covariance matrix. This is similar to a high-bandwidth approximation, but we avoid the problem of estimating centroid gains. At lower bandwidths or shorter turbulence coherence times, one needs to calculate the cross-correlation at a few more time steps, i.e. $V_{xy}(2), V_{xy}(3)$ etc in order to extrapolate by curve fitting the noise-free correlation at the origin $V'_{xy}(0)$. The noise level is then estimated from $V_{xy}(0) - V'_{xy}(0)$.

3.7 r_0 estimation

Method implemented [19, 26], undergoing validation and testing.

3.8 L_0 estimation

Method implemented, needs tuning – will try to data mine interferometer fringe tracker data and correlate with MASS/DIMM measurements and model fitting to get more insights

3.9 Residual tilt error estimation

3.10 Static aberrations

Segment aberrations and other static errors

3.11 Dynamic aberrations

Residual wind shake tip/tilt jitter and other vibrations

4 Integrated algorithm

4.1 Calibration procedures

DM and WFS biases from non-common path errors and image sharpening procedures

5 Validation and science demonstration Phase I

6 Integrated product and user interface development

7 PSF reconstruction: Phase II

Implement algorithm for LGS AO, requiring the use of a turbulence profiler. Develop algorithm for NGAO (or LTAO/MOAO/MCAO in general).

7.1 Turbulence profiling

8 Components

8.1 Anisoplanatism

The current investigation is primarily motivated by a problem in point spread function (PSF) reconstruction, where explicit knowledge of the structure function and a viable way to compute the optical transfer function (OTF) is required. Pertinent reasons for astronomers for wanting to know the PSF include the potential for more accurate photometry and astrometry, and improved deconvolution reliability when the PSF is better known, see e.g. [6, 5, 16, 8, 20, 1]. Anisoplanatism in adaptive optics has been studied in many previous instances, including comprehensive analyses by, e.g., Fried [9], Fried & Belsher [10], Sasiela [18] and Tyler [23]. These studies are chiefly concerned with calculating phase variances and Strehl ratios directly, in general not yielding intermediate results that are suitable for the problem of PSF reconstruction. In the Kolmogorov turbulence model with an infinite outer scale, the angular anisoplanatism structure function was derived by Britton [3] (see also Sect. 8.1.1). Of interest to the topic of PSF reconstruction is to know within which parameter space of outer scale and telescope size the Kolmogorov model remains a good approximation, and beyond which a von Karman model that incorporates a finite outer scale ought to be applied instead.

8.1.1 Angular anisoplanatism

Denote the optical phase of the wave-fronts integrated along the line of sight of the AO reference beam "a" and the observing direction "b" by $\phi_a(\mathbf{r},t)$ and $\phi_b(\mathbf{r},t)$, where $\mathbf{r}=(x,y)$ is a spatial coordinate in the telescope pupil plane and ϕ_b is inclined by an angle $\boldsymbol{\theta}$ to ϕ_a. The instantaneous anisoplanatism phase error ϕ_Δ is

$$\phi_\Delta(\mathbf{r},t)=\phi_a(\mathbf{r},t)-\phi_b(\mathbf{r},t), \tag{108}$$

and the anisoplanatism structure function is given by

$$\begin{aligned} D_\Delta(\mathbf{r}_1,\mathbf{r}_2,\boldsymbol{\theta}) = {} & \langle|\phi_\Delta(\mathbf{r}_1,t)-\phi_\Delta(\mathbf{r}_2,t)|^2\rangle \qquad (109)\\ = {} & \langle|\phi_a(\mathbf{r}_1)|^2\rangle+\langle|\phi_b(\mathbf{r}_1)|^2\rangle \\ & +\langle|\phi_a(\mathbf{r}_2)|^2\rangle+\langle|\phi_b(\mathbf{r}_2)|^2\rangle \\ & -2\langle\phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_1)\rangle-2\langle\phi_a(\mathbf{r}_2)\phi_b(\mathbf{r}_2)\rangle \\ & -2\langle\phi_a(\mathbf{r}_1)\phi_a(\mathbf{r}_2)\rangle-2\langle\phi_b(\mathbf{r}_1)\phi_b(\mathbf{r}_2)\rangle \\ & +2\langle\phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_2)\rangle+2\langle\phi_b(\mathbf{r}_1)\phi_a(\mathbf{r}_2).\rangle \qquad (110)\end{aligned}$$

Britton [3], based on the work of Tyler [22, 23], derives the following form for the anisoplanatism structure function under Kolmogorov turbulence statistics

$$\begin{aligned} D_\Delta(\boldsymbol{\rho},\boldsymbol{\theta}) = {} & 2.91k^2\int_0^\infty dz\, C_n^2(z) \\ & \times\left\{2|z\boldsymbol{\theta}|^{5/3}+2|\boldsymbol{\rho}|^{5/3}-|\boldsymbol{\rho}+z\boldsymbol{\theta}|^{5/3}-|\boldsymbol{\rho}-z\boldsymbol{\theta}|^{5/3}\right\}, \qquad (111)\end{aligned}$$

where $C_n^2(z)$ is the refractive index structure constant, $k=2\pi/\lambda$ and λ is the imaging wavelength. This structure function is stationary, since D_Δ depends spatially only on the separation $\boldsymbol{\rho}=\mathbf{r}_1-\mathbf{r}_2$ and not on $\mathbf{r}_1$ and $\mathbf{r}_2$ independently. To include a finite outer scale L_0, we must calculate the covariance terms in Eq. (110) using the von Karman power spectrum. Ellerbroek [7] derives a general expression for a covariance element of the type $\langle c_i c_j\rangle$ in von Karman statistics, where

$$\begin{aligned} c_i(t) &= \int d\mathbf{r}\, w_i(\mathbf{r})\,\phi_i(\mathbf{r},t), \qquad (112)\\ \phi_i(\mathbf{r},t) &= k\int_0^\infty dz\, n[\mathbf{p}_i(\mathbf{r},z),z,t], \qquad (113)\end{aligned}$$

where n is the refractive index along the ray path $\mathbf{p}_i$ of ray i, and w_i is a weighting function. In the form that applies to the current analysis, the covariance element is calculated by Ellerbroek as

$$\begin{aligned} \langle c_i c_j\rangle = {} & 0.060912k^2L_0^{5/3}\iint d\mathbf{r}\,d\mathbf{r}' w_i(\mathbf{r})w_j(\mathbf{r}') \\ & \times\int_0^\infty dz\, C_n^2(z)\, I[\alpha_{ij}(\mathbf{r},\mathbf{r}',z)]. \qquad (114)\end{aligned}$$

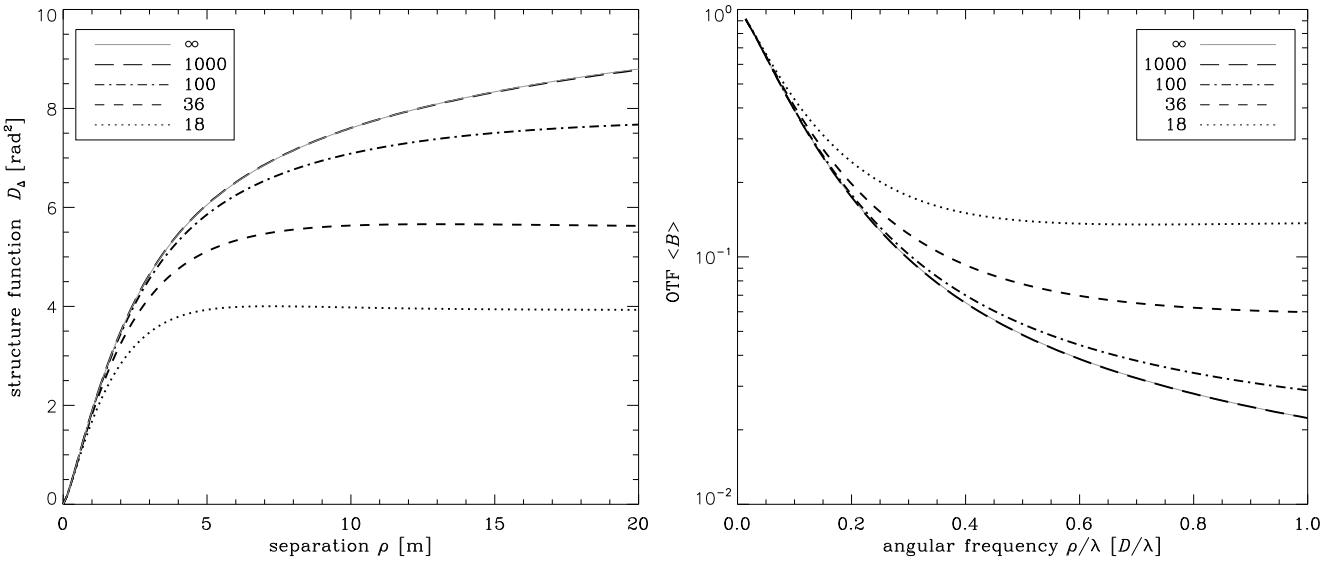

Figure 12: Left: Angular anisoplanatism (NGS) structure functions for different L_0 (cut along the positive x-axis) at $\theta = 42.4''$. Right: Angular anisoplanatism (NGS) OTFs for different L_0 (cut along the positive x-axis) at $\theta = 42.4''$.

The function $I(\alpha)$ is given by

$$I(\alpha) = \frac{\alpha^{5/6} K_{5/6}(\alpha)}{2^{5/6}\Gamma(11/6)}, \tag{115}$$

where K is a modified Bessel function of the second kind, and the argument α_{ij} is given by $\alpha_{ij} = f_0|\mathbf{p}_i(\mathbf{r}, z) - \mathbf{p}_j(\mathbf{r}', z)|$ with $f_0 = 2\pi L_0^{-1}$. Note that L_0 here denotes the propagated outer scale as observed in optical wave-fronts at the ground, and not a physical scale in the atmosphere itself. By specifying the weighting function w_i, different types of covariances can be computed. Setting $w_i(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_i)$, where δ is Dirac's delta function, turns the covariance element into the point-wise phase covariance function $\langle c_a c_b \rangle = \langle \phi_a(\mathbf{r}_1, t)\phi_b(\mathbf{r}_2, t)\rangle$, which allows us to evaluate the anisoplanatism phase structure function as a sum of variants of the general covariance function

$$\begin{aligned}\langle \phi_a(\mathbf{r}_1, t)\phi_b(\mathbf{r}_2, t)\rangle \;&=\; 0.060912 k^2 L_0^{5/3} \\ &\times \int_0^\infty dz\, C_n^2(z)\, I[\alpha_{ab}(\mathbf{r}_1, \mathbf{r}_2, z)].\end{aligned} \tag{116}$$

In ordinary NGS angular anisoplanatism, the geometry of the rays ϕ_a and ϕ_b is such that $\mathbf{p}_a(\mathbf{r}, z) = \mathbf{r}$ and $\mathbf{p}_b(\mathbf{r}, z) = \mathbf{r} - z\boldsymbol{\theta}$. Evaluating the argument α for the ten different covariance terms in Eq. (110) gives:

$$\alpha_{aa}(\mathbf{r}_1, \mathbf{r}_1) = \alpha_{bb}(\mathbf{r}_1, \mathbf{r}_1) \;=\; 0, \tag{117}$$
$$\alpha_{aa}(\mathbf{r}_2, \mathbf{r}_2) = \alpha_{bb}(\mathbf{r}_2, \mathbf{r}_2) \;=\; 0, \tag{118}$$
$$\alpha_{ab}(\mathbf{r}_1, \mathbf{r}_1) = \alpha_{ab}(\mathbf{r}_2, \mathbf{r}_2) \;=\; f_0|z\boldsymbol{\theta}|, \tag{119}$$
$$\alpha_{aa}(\mathbf{r}_1, \mathbf{r}_2) = \alpha_{bb}(\mathbf{r}_1, \mathbf{r}_2) \;=\; f_0|\mathbf{r}_1 - \mathbf{r}_2|, \tag{120}$$
$$\alpha_{ab}(\mathbf{r}_1, \mathbf{r}_2) \;=\; f_0|\mathbf{r}_1 - \mathbf{r}_2 + z\boldsymbol{\theta}|, \tag{121}$$
$$\alpha_{ba}(\mathbf{r}_1, \mathbf{r}_2) \;=\; f_0|\mathbf{r}_1 - \mathbf{r}_2 - z\boldsymbol{\theta}|. \tag{122}$$

We thus retain the property of the Kolmogorov model that the structure function is spatially stationary and only a function of $\boldsymbol{\rho} = \mathbf{r}_1 - \mathbf{r}_2$. The von Karman anisoplanatism structure function can then be expressed as

$$\begin{aligned}D_\Delta(\boldsymbol{\rho}, \boldsymbol{\theta}) \;&=\; 0.12184 k^2 L_0^{5/3} \int_0^\infty dz\, C_n^2(z) \\ &\times \{2I(0) - 2I(f_0 z\theta) - 2I(f_0\rho) \\ &+ I(f_0|\boldsymbol{\rho} + z\boldsymbol{\theta}|) + I(f_0|\boldsymbol{\rho} - z\boldsymbol{\theta}|)\},\end{aligned} \tag{123}$$

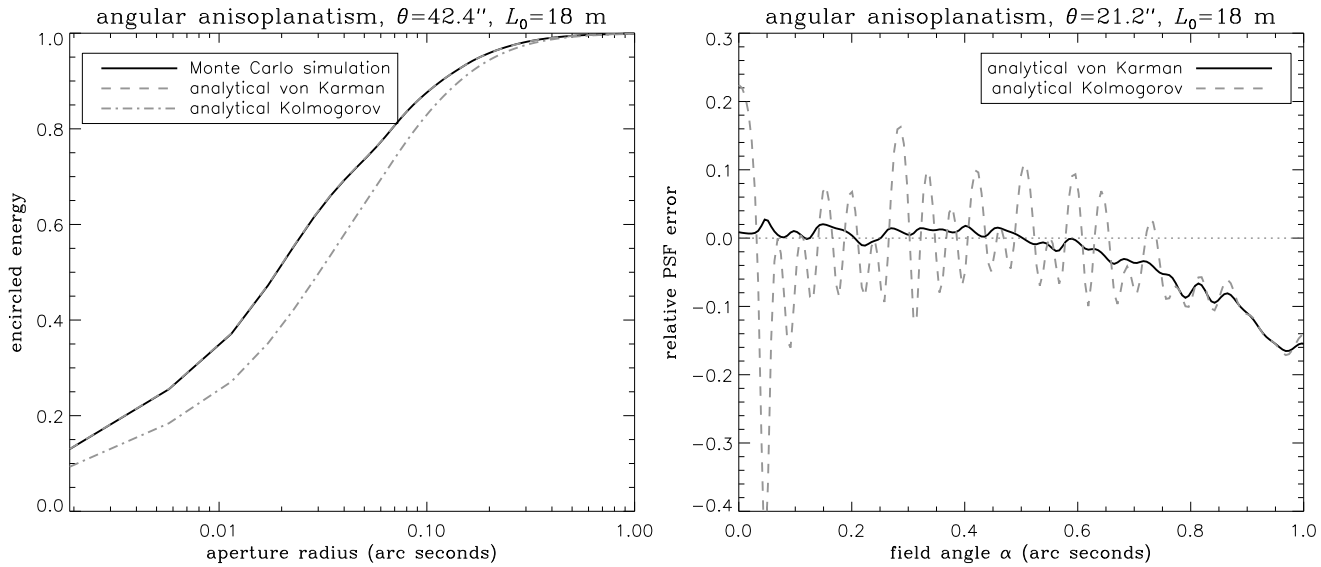

Figure 13: Left: Normalized encircled energy of PSF in the analytical von Karman and Kolmogorov models of NGS anisoplanatism, comparing to the numerical von Karman simulation. Right: Relative PSF error of the analytical von Karman (solid black line) and Kolmogorov (dashed gray line) models measured against the numerical simulation, for NGS anisoplanatism (cut of PSFs along the positive x-axis).

where $\rho = |\boldsymbol{\rho}|$, $\theta = |\boldsymbol{\theta}|$ and $I(0) = 3/5$. Since the structure function is stationary, the OTF computation required for PSF reconstruction is straightforward. In practice we calculate the expression by employing the discrete turbulence model $C_n^2(z) = \mu_0 \sum_l f_l \delta(z - z_l)$, where z_l are the layer altitudes and f_l their relative power ($\sum_l f_l = 1$), and $\mu_0 = 0.06\lambda^2 r_0^{-5/3}$. With this, the integral turns into a summation over a small number of layers, which makes the structure function readily computable.

8.1.2 Focal anisoplanatism

The structure function of focal anisoplanatism in laser guide star (LGS) AO systems can also be modeled using the methodology developed in Sect. 8.1.1. In LGS AO, the relative close proximity of the reference source requires us to treat the wave-front as a propagating spherical wave rather than a plane wave. In the geometric approximation, the result at the telescope pupil is a transversal scale change (magnification) of the wave-front, equal to $1 - z/H$, if H is the range of the LGS. The wave-front error resulting from wave-front sensing on a conical beam is called focal (or focus) anisoplanatism, or the cone-effect.

For a LGS AO system, we can derive the focal anisoplanatism structure function in the von Karman model using again the covariance element in Eq. (114). For generality we may include an additional angular offset $\boldsymbol{\theta}$, thereby combining angular and focal anisoplanatism, whereby $\mathbf{p}_a(\mathbf{r}, z) = \mathbf{r}(1 - zH^{-1})$ and $\mathbf{p}_b(\mathbf{r}, z) = \mathbf{r} - z\boldsymbol{\theta}$. This describes the intersection of a conical beam on the optical axis with a cylindrical beam offset by $\boldsymbol{\theta}$. The shorthand notation $\gamma = zH^{-1}$ will be used. The argument α_{ij} for the different covariance terms now evaluates to

$$\alpha_{ab}(\mathbf{r}_1, \mathbf{r}_1) = f_0|\gamma\mathbf{r}_1 - z\boldsymbol{\theta}|, \quad (124)$$

$$\alpha_{ab}(\mathbf{r}_2, \mathbf{r}_2) = f_0|\gamma\mathbf{r}_2 - z\boldsymbol{\theta}|, \quad (125)$$

$$\alpha_{aa}(\mathbf{r}_1, \mathbf{r}_2) = f_0|\mathbf{r}_1 - \mathbf{r}_2|(1 - \gamma), \quad (126)$$

$$\alpha_{bb}(\mathbf{r}_1, \mathbf{r}_2) = f_0|\mathbf{r}_1 - \mathbf{r}_2|, \quad (127)$$

$$\alpha_{ab}(\mathbf{r}_1, \mathbf{r}_2) = f_0|\mathbf{r}_1 - \mathbf{r}_2 - \gamma\mathbf{r}_1 + z\boldsymbol{\theta}|, \quad (128)$$

$$\alpha_{ba}(\mathbf{r}_1, \mathbf{r}_2) = f_0|\mathbf{r}_1 - \mathbf{r}_2 + \gamma\mathbf{r}_2 - z\boldsymbol{\theta}|. \quad (129)$$

We can see that the structure function will be non-stationary in this case. Changing variables according to $\mathbf{x} = \mathbf{r}_1$

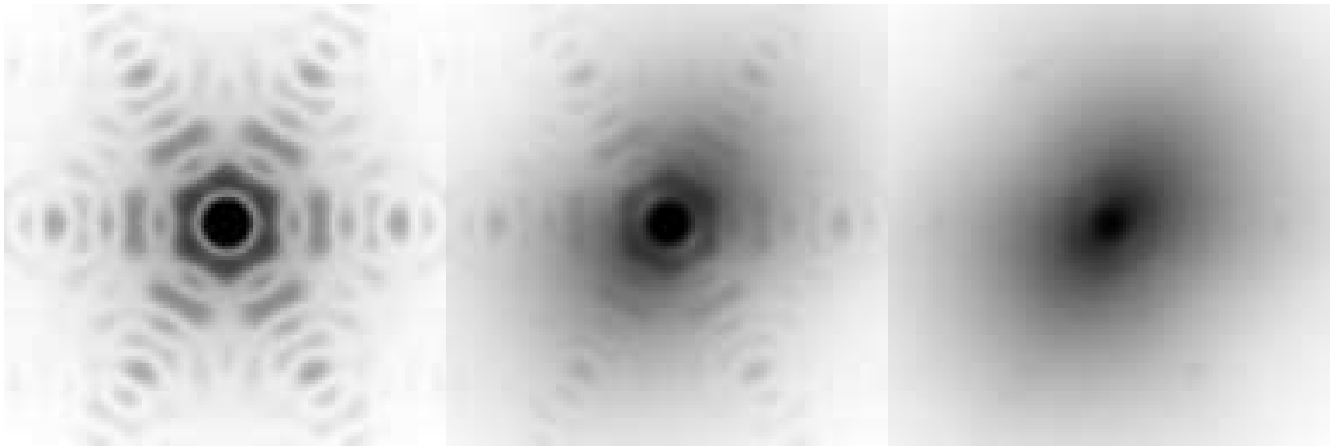

Figure 14: Focal and angular anisoplanatism PSFs (0.73×0.73 arc second field of view) in the analytical von Karman model with $L_0 = 36$ m, at $(0,0), (15,15)$ and $(30,30)$ arc seconds.

and $\boldsymbol{\rho} = \mathbf{r}_1 - \mathbf{r}_2$ gives the final result

$$\begin{aligned} D_\Delta(\mathbf{x}, \boldsymbol{\rho}, \boldsymbol{\theta}) \;=\; & 0.12184 k^2 L_0^{5/3} \int_0^H dz\, C_n^2(z) \\ & \times \{2I(0) - I[f_0\rho(1-\gamma) + I(f_0|\boldsymbol{\rho} - \gamma\mathbf{x} + z\boldsymbol{\theta}|)] \\ & - I(f_0|\gamma\mathbf{x} - z\boldsymbol{\theta}|) - I(f_0|\gamma(\boldsymbol{\rho} - \mathbf{x}) - z\boldsymbol{\theta}|) \\ & - I(f_0\rho) + I[f_0|(1-\gamma)\boldsymbol{\rho} + \gamma\mathbf{x} - z\boldsymbol{\theta}|]\}\,. \end{aligned} \tag{130}$$

This expression reduces to Eq. (123) when $\gamma = 0$. For PSF reconstruction, the non-stationary structure function does not need to be computed explicitly, only point-wise in order to evaluate the OTF integral

$$\langle B(\boldsymbol{\rho}/\lambda, \boldsymbol{\theta}) \rangle = \int d\mathbf{x}\, P(\mathbf{x}) P(\mathbf{x} + \boldsymbol{\rho}) \exp\left[-\frac{1}{2} D_\Delta(\mathbf{x}, \boldsymbol{\rho}, \boldsymbol{\theta}) \right], \tag{131}$$

where P is the aperture transmission function and $\langle B \rangle$ is the long-exposure OTF. The validity of Eq. (131) rests on the assumptions that there is no scintillation and that the residual phase ϕ_Δ follows Gaussian statistics. A common technique for computing the OTF from a non-stationary structure function is to invoke the pupil-averaging approximation, as introduced by Véran [28]. Advances in computing power have partly rendered this approximation unnecessary, and the focal anisoplanatism OTF can be readily computed point-wise using the exact expressions in Eq. (130) and Eq. (131).

8.1.3 Sample numerical results

Figures 12-16 present some sample numerical results, comparing the analytical von Karman model in Eq. (123) and Eq. (130) to numerical Monte Carlo simulations, and in the case of NGS angular anisoplanatism also comparing to the analytical Kolmogorov model in Eq. (111). For all numerical results, the W.M. Keck Observatory (WMKO) telescope and AO system was used as the model, which gives the characteristic PSF shape shown in Fig. 14. The turbulence model was the 7-layer Mauna Kea Ridge (MKR) model[1] with an overall $r_0 = 0.16$ m, $\theta_0 = 2.57''$ and $d_0 = 4.65$ m (for a definition of these quantities see e.g. [14]). All calculations were done at 2 μm imaging wavelength and at zenith, and the LGS beacon altitude was 90 km. The PSFs were computed at three evaluation points $(\theta_x, \theta_y) = \{(0,0), (15,15), (30,30)\}$, equaling radial off-axis offsets of, respectively, 0, 21.2 and 42.4 arc seconds. The outer scale in the analytical von Karman calculations varied between $L_0 = \{1000, 100, 36, 18\}$ meters, while the

[1]The MKR model was produced within the framework of the WMKO Next-Generation Adaptive Optics (NGAO) project in collaboration with the Thirty-Meter Telescope (TMT) site monitoring campaign, by use of their MASS/DIMM instrument on Mauna Kea. The statistical data are currently proprietary of TMT and are therefore not reproduced in this paper.

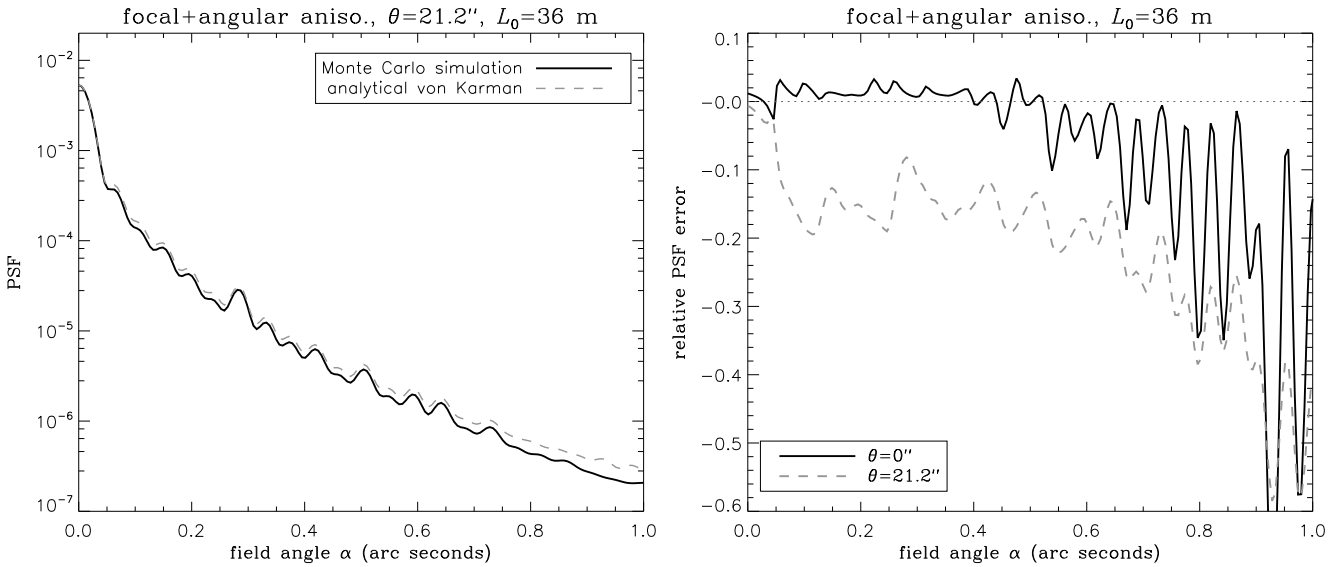

Figure 15: Left: Sample LGS anisoplanatism PSF profile (cut along the positive x-axis), comparing the numerical simulation (solid black line) and the analytical von Karman model. Right: Relative PSF error of the analytical von Karman model for LGS anisoplanatism, measured against the numerical simulation for the $\theta = 0''$ and $\theta = 21.2''$ cases (cut of PSFs along the positive x-axis).

numerical simulation only produced results at $L_0 = \{36, 18\}$ meters (since the phase screens were FFT-based with a length of 72 meters).

An overview of Strehl ratios obtained in the various cases are given in Table 1, showing exancellent agreement between the analytical von Karman model and numerical simulations. The von Karman model also reproduces the Kolmogorov result at very large outer scales. The effect of the finite outer scale on the structure function and the OTF can be plotted in the case of NGS anisoplanatism (since it is stationary), as shown in Figs. 12 and 12. Figures 13 and 15 plot the relative PSF error, i.e. the normalized PSF subtraction computed as (numerical − analytical)/numerical, for a line segment of the PSFs (the positive x-axis). While in some parts the reported numbers may seem large (0.5 implies a 50% error), the relative error in the PSF halo should be weighted by the fact that the energy levels are two or three orders of magnitude lower than in the central region of the PSF. The encircled energy curves in Fig. 13 and 16 also show that even when the relative error in the wings of the PSF is large, this still amounts to a negligible error in an absolute sense. As long as the overall energy distribution is correct, a small error in the PSF halo has no impact on the encircled energy. Even in the worst case considered (dashed gray curves in Fig. 15 and 15), the analytical von Karman model is still very accurate and the error practically negligible.

8.1.4 Spatial filtering

In a real AO system, anisoplanatism manifests only on the subspace of controlled modes. Applying the full power spectrum of turbulence will overestimate the anisoplanatism error by counting the fitting error to the anisoplanatism error budget. If the fitting error is small, this approximation might be acceptable. In order to project the anisoplanatism wave-front error onto controlled modes we rewrite Eq. (108) as

$$\phi_\Delta(\mathbf{r}, t) = \phi_{\|a}(\mathbf{r}, t) - \phi_{\|b}(\mathbf{r}, t), \tag{132}$$

where the notation symbolizes that we have retained only the low-frequency (controlled) part $\phi_\|$ of the decomposition $\phi = \phi_\| + \phi_\perp$. This splitting of the phase into two orthogonal components in spatial frequency domain is common in AO PSF modeling and PSF reconstruction methodology, see e.g. [17, 28, 15, 11]. In practice the domain of $\phi_\|$ is defined as the vector space spanned by the set of N_a influence functions $\{h_i(\mathbf{r})\}_{i=1}^{N_a}$ of the DM. Defining w_i as the

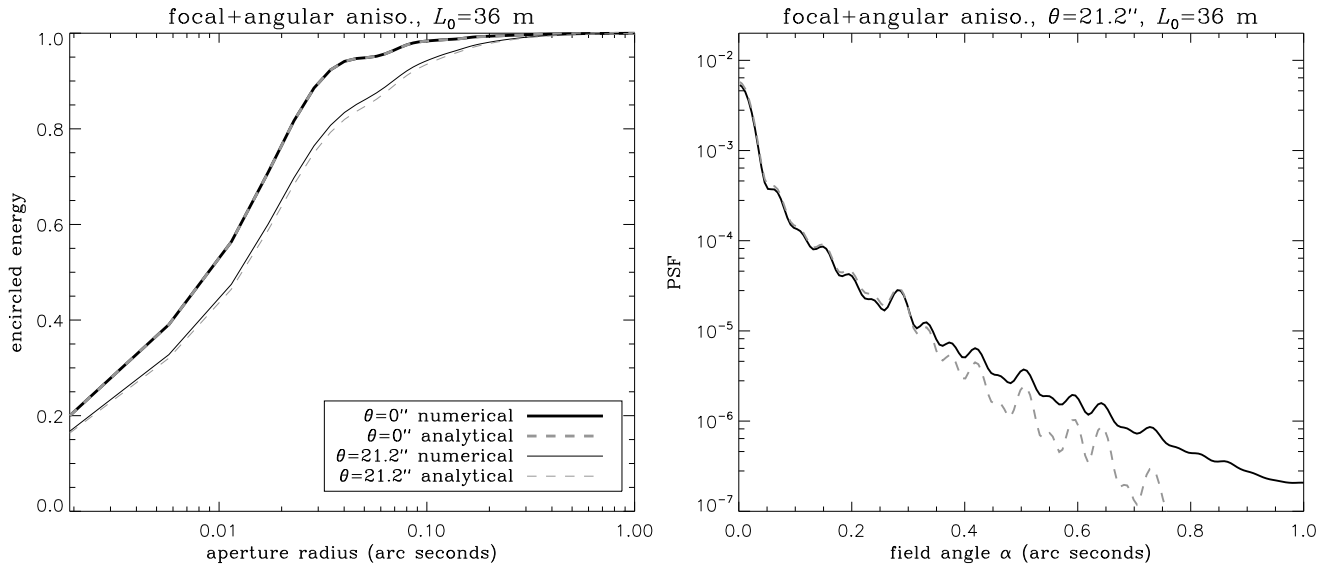

Figure 16: Left: Normalized encircled energy of PSF in the analytical von Karman model of LGS anisoplanatism, comparing to the numerical von Karman simulation. Right: Simulation showing the effect of spatial filtering on LGS anisoplanatism.

piston-removed influence function, we have the relations

$$\phi_{\|a}(\mathbf{r},t) = \sum_{i=i}^{N_a} a_i(t) w_i(\mathbf{r}), \tag{133}$$

$$a_i(t) = \int d\mathbf{r}\, w_i(\mathbf{r}) \phi_a(\mathbf{r},t), \tag{134}$$

$$w_i(\mathbf{r}) = P(\mathbf{r})[h_i(\mathbf{r}) - p_i], \tag{135}$$

$$p_i = \frac{\int d\mathbf{r}\, P(\mathbf{r})\, h_i(\mathbf{r})}{\int d\mathbf{r}\, P(\mathbf{r})}, \tag{136}$$

where p_i is the piston of each mode and $\{a_i\}_{i=1}^{N_a}$ is a set of expansion coefficients (i.e. actuator commands). Analogously, the off-axis beam ϕ_b is projected onto the set $\{b_i\}_{i=1}^{N_a}$. The covariance function now has the general form

$$\langle \phi_{\|a}(\mathbf{r}_1) \phi_{\|b}(\mathbf{r}_2) \rangle = \sum_{i=1}^{N_a} \sum_{j=1}^{N_a} \langle a_i b_j \rangle w_i(\mathbf{r}_1) w_j(\mathbf{r}_2). \tag{137}$$

The covariance matrix $\langle a_i b_j \rangle$ is calculated from Eq. (114). When Eq. (133) is substituted into the expression for the structure function, ten terms of the form above result, containing the two covariance matrices $A_{ij} = \langle a_i a_j \rangle$ and $B_{ij} = \langle a_i b_j \rangle$. Defining the function $V_{ij}(\boldsymbol{\theta} z_l)$ as

$$V_{ij}(\boldsymbol{\theta} z_l) = 0.144 \left(\frac{L_0}{r_0}\right)^{5/3} \int d\mathbf{r}\, w_i(\mathbf{r}) [I * w_j](\mathbf{r} + \boldsymbol{\theta} z_l), \tag{138}$$

where asterisk ($*$) denotes a two-dimensional convolution and I is the function defined in Eq. (115), we have that

$$A_{ij} = V_{ij}(0), \qquad B_{ij} = \sum_{l=1}^{N_l} f_l V_{ij}(\boldsymbol{\theta} z_l). \tag{139}$$

Collecting and factoring terms, the structure function can then be expressed as

$$D_\Delta(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\theta}) = \sum_{i=1}^{N_a} \sum_{j=1}^{N_a} (2A_{ij} - B_{ij} - B_{ji}) \times [w_i(\mathbf{r}_1) - w_i(\mathbf{r}_2)] [w_j(\mathbf{r}_1) - w_j(\mathbf{r}_2)]. \tag{140}$$

Introducing the notation Q for the field-dependent symmetric matrix $Q = 2A - B - B^T$ and performing the variable substitution $\{\mathbf{x} = \mathbf{r}_2, \boldsymbol{\rho} = \mathbf{r}_1 - \mathbf{r}_2\}$ gives

$$D_\Delta(\mathbf{x}, \boldsymbol{\rho}, \boldsymbol{\theta}) = \sum_{i=1}^{N_a} \sum_{j=1}^{N_a} Q_{ij} \left[w_i(\mathbf{x}) - w_i(\mathbf{x} + \boldsymbol{\rho})\right] \times \left[w_j(\mathbf{x}) - w_j(\mathbf{x} + \boldsymbol{\rho})\right]. \tag{141}$$

Readers familiar with PSF reconstruction techniques will recognize this form, as it appears in e.g. [28] before the pupil-averaging approximation is applied. Using the results of Gendron et al. [11] we can compute the OTF without invoking the pupil-averaging approximation, however. By diagonalizing the covariance matrix $Q = S\Lambda S^T$, with $\{\sigma_k\}_{k=1}^{N_a}$ being the singular values of the diagonal matrix $\Lambda_{ik} = \delta_{ik}\sigma_k$, we obtain the transformed modal basis

$$h'_k(\mathbf{x}) = \sum_{i=1}^{N_a} S_{ik} w_i(\mathbf{x}). \tag{142}$$

In this basis, the structure function can be expressed in the computationally tractable form

$$D_\Delta(\mathbf{x}, \boldsymbol{\rho}, \boldsymbol{\theta}) = \sum_{k=1}^{N_a} \sigma_k \left|h'_k(\mathbf{x}) - h'_k(\mathbf{x} + \boldsymbol{\rho})\right|^2, \tag{143}$$

which enables the OTF to be calculated efficiently. This method is computationally viable to be employed for PSF reconstruction from non-stationary structure functions in general, and not just anisoplanatism, so long as the structure functions can be expressed on the general form in Eq. (141).

8.1.5 Conclusions

Structure functions for angular and focal anisoplanatism were derived analytically including a finite outer scale in a von Karman type turbulence model. The analytical result is in excellent agreement with numerical Monte Carlo simulations, and offers an analytical method to produce anisoplanatism kernels for PSF reconstruction problems and PSF modeling in general. The model may replace Kolmogorov-based models where the impact of a finite outer scale becomes non-negligible. In the current investigation it is found that for 10-meter-class telescopes, an outer scale smaller than ∼30 meters may have a non-negligible impact on the anisoplanatism estimation, and under these conditions the von Karman model should be preferred.

Not taken into account in the current analysis is the spatial filtering of the wave-front performed by the AO system. This effect can be accounted for by setting the weighting functions w_i equal to the deformable mirror (DM) influence functions in Eq. (114). This renders the structure function non-stationary, but in the case of NGS anisoplanatism the expression can still be put on a form that allows for a relatively efficient computation of the OTF. For completeness, this result is given in Appendix A, but the method is unlikely to be employed for PSF reconstruction since: 1) in the case of focal anisoplanatism, the expression becomes computationally expensive and impractical for PSF reconstruction, so the approximation in Eq. (130) must be invoked in any case, and 2) the effect is small in high-order AO systems that have $d/r_0 < 1$, where d is the DM actuator spacing. As shown in Fig. 16, the effect of spatial filtering is to reduce the amount of scattered light from anisoplanatism in the PSF halo, which should be modeled by the fitting error instead. If the fitting error is small, however, omitting spatial filtering in the anisoplanatism model is a small effect on the the PSF structure, on the Strehl ratio, and on the resulting photometry estimates produced with this model. For the current WMKO AO system ($d = 0.56$ m), in this model the LGS Strehl ratio at $L_0 = 36$ m at the three evaluations points without spatial filtering is $[0.792, 0.384, 0.134]$, and with spatial filtering they are $[0.829, 0.416, 0.0.147]$. Given that the added complexity of trying to model this effect realistically makes it impractical for LGS PSF reconstruction, one possible compromise might be to use the unfiltered models derived in Sect. 8.1.1 and Sect. 8.1.2, and implement a heuristic adjustment to the fitting error in order to account for the fact that some high-spatial-frequency wavefront error is added by the anisoplanatism term. This route is being investigated further within the PSF reconstruction project that motivated the current research.

$\theta =$	$0''$	$21.2''$	$42.4''$
NGS ANISOPLANATISM			
analytical Kolmogorov			
$L_0 = \infty$	1.000	0.343	0.101
analytical von Karman			
$L_0 = 1000$	1.000	0.343	0.101
$L_0 = 100$	1.000	0.350	0.106
$L_0 = 36$	1.000	0.380	0.127
$L_0 = 18$	1.000	0.438	0.183
numerical von Karman			
$L_0 = 36$	1.000	0.383	0.127
$L_0 = 18$	1.000	0.442	0.183
LGS ANISOPLANATISM			
analytical von Karman			
$L_0 = 36$	0.782	0.387	0.143
$L_0 = 18$	0.803	0.440	0.192
numerical von Karman			
$L_0 = 36$	0.791	0.385	0.137
$L_0 = 18$	0.812	0.445	0.192

Table 1: NGS and LGS anisoplanatism Strehl ratios, comparing analytical predictions to numerical simulations.

8.2 Tilt anisoplanatism with TT NGS

8.3 Tomography error with multiple LGS

8.4 Tilt anisoplanatism with multiple NGS

This part is well developed already (see analysis on NGAO null-modes)

8.5 List of assumptions, approximations and uncertainties

1. DM model (51) linearity: approximation (fairly good)
2. $\langle\widetilde{\phi}(f)\widetilde{\phi}(f')\rangle = \delta(f-f')\langle\widetilde{\phi}(f)\widetilde{\phi}(f')\rangle$ - standard Kolmogorov assumption, spatial frequencies are uncorrelated
3. Uniform field amplitude (no scintillation)
4. Gaussian statistics for the residual phase error
5. Parallel and orthogonal phase components are statistically uncorrelated
6. Structure function of the residual phase is homogeneous: not really true

9 Validation and science demonstration Phase II

References

[1] M. Aubailly, M. C. Roggemann, and T. J. Schulz. Approach for reconstructing anisoplanatic adaptive optics images. *Appl. Opt.*, 46:6055–6063, August 2007.

[2] M. Born and E. Wolf. *Principles of optics*. Pergamon Press, Oxford, 6th edition, 1980.

[3] M. C. Britton. The Anisoplanatic Point-Spread Function in Adaptive Optics. *Publ. Astr. Soc. Pac.*, 118:885–900, June 2006.

[4] A. Consortini and L. Ronchi. Choice of the model of atmospheric turbulence. *Appl. Opt.*, 11:1205–+, May 1972.

[5] G. Cresci, R. I. Davies, A. J. Baker, and M. D. Lehnert. Accounting for the anisoplanatic point spread function in deep wide-field adaptive optics images. *Astron. Astrophys.*, 438:757–767, August 2005.

[6] E. Diolaiti, O. Bendinelli, D. Bonaccini, L. M. Close, D. G. Currie, and G. Parmeggiani. Starfinder: an idl gui based code to analyze crowded fields with isoplanatic correcting psf fitting. In Peter L. Wizinowich, editor, *Adaptive Optical Systems Technology*, volume 4007 of *Proc. SPIE*, pages 879–887, 2000.

[7] B. L. Ellerbroek. Including outer scale effects in zonal adaptive optics calculations. *Appl. Opt.*, 36:9456–9567, 1997.

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

[9] D. L. Fried. Anisoplanatism in adaptive optics. *J. Opt. Soc. Am.*, 72(1):52–61, 1982.

[10] D. L. Fried and J. F. Belsher. Analysis of fundamental limits to artificial-guide-star adaptive-optics-system performance for astronomical imaging. *J. Opt. Soc. Am. A*, 11(1):277–287, January 1994.

[11] E. Gendron, Y. Clénet, T. Fusco, and G. Rousset. New algorithms for adaptive optics point-spread function reconstruction. *Astron. Astrophys.*, 457:359–363, October 2006.

[12] E. Gendron and P. Léna. Astronomical adaptive optics I. Modal control optimization. *Astron. Astrophys.*, 291:337–347, 1994.

[13] E. Gendron and P. Léna. Astronomical adaptive optics II. Experimental results of an optimized modal control. *Astron. Astrophys. Suppl. Ser.*, 111:153–167, 1995.

[14] J. W. Hardy. *Adaptive optics for Astronomical Telescope*. Oxford University Press, 1998.

[15] L. Jolissaint, J.-P. Veran, and J. Marino. OPERA, an automatic PSF reconstruction software for Shack-Hartmann AO systems: application to Altair. In D. Bonaccini Calia, B. L. Ellerbroek, and R. Ragazzoni, editors, *Advancements in Adaptive Optics. Edited by Domenico B. Calia, Brent L. Ellerbroek, and Roberto Ragazzoni. Proceedings of the SPIE, Volume 5490, pp. 151-163 (2004).*, volume 5490 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 151–163, October 2004.

[16] Y. Minowa, N. Kobayashi, Y. Yoshii, T. Totani, T. Maihara, F. Iwamuro, H. Takami, N. Takato, Y. Hayano, H. Terada, S. Oya, M. Iye, and A. T. Tokunaga. Subaru Super Deep Field with Adaptive Optics. I. Observations and First Implications. *Astrophys. J.*, 629:29–44, August 2005.

[17] F. J. Rigaut, J. Veran, and O. Lai. Analytical model for Shack-Hartmann-based adaptive optics systems. In Domenico Bonaccini and Robert K. Tyson, editors, *Adaptive Optical System Technologies*, volume 3353 of *Proc. SPIE*, pages 1038–1048, September 1998.

[18] R. J. Sasiela. Strehl ratios with various types of anisoplanatism. *J. Opt. Soc. Am. A*, 9(8):1398–1406, August 1992.

[19] M. Schöck, D. Le Mignant, G. A. Chanan, P. L. Wizinowich, and M. A. van Dam. Atmospheric turbulence characterization with the Keck adaptive optics systems. I. Open-loop data. *Appl. Opt.*, 42:3705–3720, July 2003.

[20] C. D. Sheehy, N. McCrady, and J. R. Graham. Constraining the Adaptive Optics Point-Spread Function in Crowded Fields: Measuring Photometric Aperture Corrections. *Astrophys. J.*, 647:1517–1530, August 2006.

[21] A. Tokovinin. From Differential Image Motion to Seeing. *Publ. Astr. Soc. Pac.*, 114:1156–1166, October 2002.

[22] G. A. Tyler. Merging: a new method for tomography through random media. *J. Opt. Soc. Am. A*, 11:409–424, January 1994.

[23] G. A. Tyler. Wave-front compensation for imaging with off-axis guide stars. *J. Opt. Soc. Am. A*, 11:339–346, January 1994.

[24] G. A. Tyler. Reconstruction and assessment of the least-squares and slope discrepancy components of the phase. *Journal of the Optical Society of America A*, 17:1828–1839, October 2000.

[25] M. A. van Dam. Measuring the centroid gain of a Shack-Hartmann quad-cell wavefront sensor by using slope discrepancy. *Journal of the Optical Society of America A*, 22:1509–1514, August 2005.

[26] M. A. van Dam, A. H. Bouchez, D. Le Mignant, E. M. Johansson, P. L. Wizinowich, R. D. Campbell, J. C. Y. Chin, S. K. Hartman, R. E. Lafon, P. J. Stomski, Jr., and D. M. Summers. The W. M. Keck Observatory Laser Guide Star Adaptive Optics System: Performance Characterization. *Publ. Astr. Soc. Pac.*, 118:310–318, February 2006.

[27] M. A. van Dam, D. Le Mignant, and B. A. Macintosh. Performance of the Keck Observatory Adaptive-Optics System. *Appl. Opt.*, 43:5458–5467, October 2004.

[28] J.-P. Véran, F. Rigaut, H. Maître, and D. Rouan. Estimation of the adaptive optics long exposure point spread function using control loop data. *J. Opt. Soc. Am. A*, 14(11), 1997.

[29] Jean-Pierre Véran. *Estimation de la réponse impulsionnelle et restauration d'image en optique adaptative. Application au système d'optique adaptative du Télescope Canada-France-Hawaii.* PhD thesis, Ecole Nationale Supérieure des Télécommunication, Paris, Observatiore de Paris-Meudon, 1997.