Lecture 12

Overall characterization and analysis of an AO system
(Work in progress)

Harrison H. Barrett
University of Arizona

Overall characterization and analysis of an AO system

- In Lecture 10 we showed how to apply statistical estimation theory to a wavefront sensor. Stochastic effects included photon noise and readout noise in the sensor but not the atmosphere, which was assumed to be frozen.

- In Lecture 11 we ignored the wavefront sensor and looked at the effect of the deformable mirror. The atmosphere was treated stochastically,

- Goal today is to look at the big picture and show how to do a full deterministic and stochastic characterization of an AO system – and to relate the characterization to task performance.

OUTLINE

- (Skip the review today)
- Conventional characterizations – and why they are inappropriate!
- What do we really want to know – and why?
- Deterministic and stochastic effects to consider
- Multiply stochastic averaging – basics and an example
- Triply stochastic averaging in AO
- Questions for the group

References

Barrett and Myers, FOUNDATIONS OF IMAGE SCIENCE

Chap. 7, Deterministic Descriptions of Imaging Systems

Chap. 8, Stochastic Descriptions of Objects and Images

Chnaps. 16-19 in B&M give detailed examples of how to apply the methods of Chaps. 7 and 8 to specific imaging systems.

Today's lecture can be regarded as a start on a "Chap. 20", Application to Adaptive Optics.

Additional references: The OAIQs

OAIQ I: HHB, "Objective assessment of image quality: effects of quantum noise and object variability", J. Opt. Soc. Am. A, 7:1266-1278, 1990.

OAIQ II: HHB et al., "Objective assessment of image quality: II. Fisher information, Fourier crosstalk, and figures of merit for task performance", J. Opt. Soc. Am. A, 12, 5:834-852, 1995.

OAIQ III: HHB et al., "Objective assessment of image quality: III. ROC metrics, ideal observers and likelihood-generating functions", J. Opt. Soc. Am. A, 15:1520-1535, 1998.

Today's lecture can be regarded as a start on an OAIQ IV, Application to Adaptive Optics.

Conventional characterizations of imaging systems

- Deterministic properties
  - Sensitivity, collection efficiency, étendue, quantum efficiency, ...
  - Point spread function (PSF)
  - Modulation transfer function (MTF)
- Stochastic properties
  - Image variance
  - Signal-to-noise ratio, contrast-to-noise ratio
  - Noise power spectrum (NPS)
  - Detective quantum efficiency (DQE $=$ MTF^2 / NPS)

Problems with conventional characterizations

- Real systems have discrete outputs
- Systems are never shift-invariant
- Noise is not stationary (even in a discrete sense)
- Noise often depends on the object (Poisson statistics)
- At best, they are just characterizations, not performance measures

More general characterizations

- Deterministic properties

  - Conditional mean of the image data
    Conditional on a specific object, $\overline{\mathbf{g}}(\mathbf{f})$
    Conditional on a state of nature or hypothesis (e.g., signal absent)

  - Sensitivity function $h_m(\mathbf{r})$ (for linear system): $\overline{g}_m(\mathbf{f}) = \int d^2r \, h_m(\mathbf{r}) \, f(\mathbf{r})$

- Stochastic properties

  - Conditional PDF of the data
    Conditional on a specific object, $\mathrm{pr}(\mathbf{g}|\mathbf{f})$
    Conditional on a state of nature or hypothesis, $\mathrm{pr}(\mathbf{g}|H_j)$

  - Conditional covariance matrices

Relation of general and conventional descriptions

Sensitivity function is a general PSF for linear systems. To get from $h_m(\mathbf{r})$ to usual $p(\mathbf{r})$, must assume:

- Infinite amount of data (infinitesimal detector pixels)
- Shift invariance (no borders, vignetting, $1/r^2$, anisoplanatism, ...)

To use NPS, must assume:

- Stationary noise (no Poisson effects, infinite detector field, ...)

To use DQE, must assume:

- Linear, shift-invariant system
- Stationary, Gaussian noise

What do we really, really want to know about the system?

Short answer: Task performance

Slightly longer answer: Properties that let us compute task performance

Minimal set:

- Conditional mean image (or h_m if system linear)
- Conditional covariance matrix

Definition of conditional covariance

$$\left[\mathbf{K}_{\mathbf{g}|\mathbf{f}}\right]_{mm'} \equiv \langle [g_m - \overline{g}_m] \; [g_{m'} - \overline{g}_{m'}] \rangle_{\mathbf{g}|\mathbf{f}}$$

Specifies how image values at pixels m and m' covary.

$\boxed{\text{Why covariance? (Isn't the variance enough?)}}$

Answer 1: Correlated noise doesn't average out

Suppose pixel variance is σ^2 but object of interest covers N pixels. If noise is uncorrelated:

Total signal $\propto N$

Total noise variance $\propto N$

SNR $=$ Total signal $/\sqrt{\text{Total variance}} \propto \sqrt{N}$

But this gain is not obtained with correlated noise.

$\boxed{\text{Why covariance?}}$

Answer 2: Detection or discrimination tasks

Consider two hypotheses:

H_0: $\langle \mathbf{g} | H_0 \rangle \equiv \overline{\mathbf{g}}_0$, Covariance matrix $\mathbf{K}_0$

H_1: $\langle \mathbf{g} | H_1 \rangle \equiv \overline{\mathbf{g}}_1$, Covariance matrix $\mathbf{K}_1$

Optimum linear discriminant (Hotelling):

$$t(\mathbf{g}) = [\overline{\mathbf{g}}_1 - \overline{\mathbf{g}}_0]^t \mathbf{K}^{-1} \mathbf{g} \,, \qquad \mathbf{K} = \tfrac{1}{2}\mathbf{K}_0 + \tfrac{1}{2}\mathbf{K}_1$$

Performance measure (Hotelling SNR^2 or Hotelling trace):

$$\mathrm{SNR}^2 = [\overline{\mathbf{g}}_1 - \overline{\mathbf{g}}_0]^t \mathbf{K}^{-1} [\overline{\mathbf{g}}_1 - \overline{\mathbf{g}}_0]$$

Many variants: Can use mean and covariance of raw images, reconstructions, channel outputs, features, etc. Can apply to linear or nonlinear systems.

Don't need (and usually can't use) PSF, MTF, NPS, DQE, ...

Why covariance?

Answer 3: Estimation tasks

Suppose you want to estimate a scalar parameter of the form,

$$\theta = \boldsymbol{\chi}^\dagger \mathbf{f} = \int d^2r\, \chi^*(\mathbf{r})\, f(\mathbf{r})\,,$$

from data satisfying

$$\mathbf{g} = \boldsymbol{\mathcal{H}}\mathbf{f} + \mathbf{n}\,.$$

The minimum-variance unbiased (Gauss-Markov) estimator is

$$\hat{\theta} = \boldsymbol{\chi}^\dagger \left(\boldsymbol{\mathcal{H}}^\dagger \mathbf{K}_{\mathbf{g}}^{-1} \boldsymbol{\mathcal{H}}\right)^{-1} \boldsymbol{\mathcal{H}}^\dagger \mathbf{K}_{\mathbf{g}}^{-1} \mathbf{g}\,,$$

and the resulting variance is

$$\mathrm{Var}\{\hat{\theta}\} = \boldsymbol{\chi}^\dagger \left(\boldsymbol{\mathcal{H}}^\dagger \mathbf{K}_{\mathbf{g}}^{-1} \boldsymbol{\mathcal{H}}\right)^{-1} \boldsymbol{\chi}\,.$$

Again, many variants possible, but all require knowledge of $\boldsymbol{\mathcal{H}}$ and $\mathbf{K}_{\mathbf{g}}$.

$\boxed{\text{Summary so far}}$

Good news: For many choices of task and observer, figure of merit can be computed from knowledge of mean image vector and covariance matrix.

Bad news: You gotta invert the covariance

Good news: There are lots of neat tricks for the inversion

Bad news: I don't have time to discuss the tricks today

Good news: What I will discuss today is how to compute the covariance matrix in the first place.

$\boxed{\text{Random effects}}$

In conventional digital imaging:

- Photon noise and electronic noise
- Random object (e.g., cluttered background)

In adaptive optics:

- All of above, plus:
- Random atmosphere
- Photon and electronic noise in WFS
- Resulting random PSF

Methodology for systematically including all of these effects:
Multiply stochastic averaging.

Multiply stochastic averaging: An example

Linear CD imaging system, Poisson noise

Conditional mean image (averaged over Poisson noise for given object):

$$\overline{g}_m = \int d^2r \, h_m(\mathbf{r}) f(\mathbf{r}) \,, \qquad \mathsf{Pr}(\mathbf{g}|\mathbf{f}) = \prod_{m=1}^{M} \mathsf{exp}(-\overline{g}_m) \frac{[\overline{g}_m]^{g_m}}{g_m!} \,.$$

The conditional covariance matrix is given in component form as

$$[\mathbf{K}_{\mathbf{g}|\mathbf{f}}]_{mm'} = \langle [g_m - \overline{g}_m] \; [g_{m'} - \overline{g}_{m'}] \rangle_{\mathbf{g}|\mathbf{f}} = \overline{g}_m \, \delta_{mm'} \,,$$

or in outer-product form as

$$\mathbf{K}_{\mathbf{g}|\mathbf{f}} = \left\langle [\Delta \mathbf{g}] \, [\Delta \mathbf{g}]^t \right\rangle_{\mathbf{g}|\mathbf{f}} = \mathsf{diag} \; (\overline{\mathbf{g}}) \,, \qquad \Delta \mathbf{g} \equiv \mathbf{g} - \overline{\mathbf{g}} \,.$$

Note that $\mathbf{K}_{\mathbf{g}|\mathbf{f}}$ is a function of $\mathbf{f}$ (through $\overline{\mathbf{g}}$).

Continuation of the example – averaging over $\mathbf{f}$

The overall mean image (averaged over both Poisson noise and object randomness) is defined by

$$\overline{\overline{g}}_m = \langle g_m \rangle_{\mathbf{g},\mathbf{f}} = \int d\mathbf{g} \int d\mathbf{f}\; g_m \;\mathrm{pr}(\mathbf{g},\mathbf{f}) = \int d\mathbf{f} \int d\mathbf{g}\, g_m \;\mathrm{pr}(\mathbf{g}|\mathbf{f})\; \mathrm{pr}(\mathbf{f})\,.$$

Operationally, all we do is throw on another bar:

$$\overline{\overline{g}}_m = \langle \overline{g}_m \rangle_{\mathbf{f}} = \int d^2 r\, h_m(\mathbf{r}) \overline{f}(\mathbf{r})\,.$$

The overall covariance matrix is defined by

$$\mathbf{K}_{\mathbf{g}} \equiv \left\langle [\mathbf{g} - \overline{\overline{\mathbf{g}}}][\mathbf{g} - \overline{\overline{\mathbf{g}}}]^t \right\rangle_{\mathbf{g},\mathbf{f}} = \left\langle \left\langle [\mathbf{g} - \overline{\overline{\mathbf{g}}}][\mathbf{g} - \overline{\overline{\mathbf{g}}}]^t \right\rangle_{\mathbf{g}|\mathbf{f}} \right\rangle_{\mathbf{f}}\,.$$

Continuation of the example – decomposition of the covariance matrix

From last slide:

$$\mathbf{K_g} \equiv \langle [\mathbf{g} - \overline{\overline{\mathbf{g}}}][\mathbf{g} - \overline{\overline{\mathbf{g}}}]^t \rangle_{\mathbf{g},\mathbf{f}} = \langle \langle [\mathbf{g} - \overline{\overline{\mathbf{g}}}][\mathbf{g} - \overline{\overline{\mathbf{g}}}]^t \rangle_{\mathbf{g}|\mathbf{f}} \rangle_{\mathbf{f}} \,.$$

Now add and subtract $\overline{\mathbf{g}}$ in each factor:

$$\mathbf{K_g} = \langle \langle [\mathbf{g} - \overline{\mathbf{g}} + \overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}}][\mathbf{g} - \overline{\mathbf{g}} + \overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}}]^t \rangle_{\mathbf{g}|\mathbf{f}} \rangle_{\mathbf{f}}$$

$$= \langle \langle [\mathbf{g} - \overline{\mathbf{g}}][\mathbf{g} - \overline{\mathbf{g}}]^t \rangle_{\mathbf{g}|\mathbf{f}} \rangle_{\mathbf{f}} + \langle [\overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}}]\,[\overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}}]^t \rangle_{\mathbf{f}}$$

Note that the cross term has vanished identically since

$$\langle \langle [\mathbf{g} - \overline{\mathbf{g}}][\overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}}]^t \rangle_{\mathbf{g}|\mathbf{f}} \rangle_{\mathbf{f}} = \langle \langle [\mathbf{g} - \overline{\mathbf{g}}] \rangle_{\mathbf{g}|\mathbf{f}} [\overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}}]^t \rangle_{\mathbf{f}} = 0$$

Thus, with no assumptions about independence of $\mathbf{g}$ and $\mathbf{f}$, we can write

$$\mathbf{K_g} = \overline{\mathbf{K}}_{\mathbf{g}}^{Pois} + \mathbf{K}_{\overline{\mathbf{g}}}^{obj}$$

where the first term is a diagonal matrix representing the Poisson noise, and the second term arises from object variability.

More on the covariance decomposition

$$\mathbf{K_g} = \overline{\mathbf{K}}_{\mathbf{g}}^{Pois} + \mathbf{K}_{\overline{\mathbf{g}}}^{obj}$$

Look at first term in more detail:

$$\overline{\mathbf{K}}_{\mathbf{g}}^{Pois} = \langle\langle[\mathbf{g} - \overline{\mathbf{g}}\,][\mathbf{g} - \overline{\mathbf{g}}\,]^t\,\rangle_{\mathbf{g}|\mathbf{f}}\rangle_{\mathbf{f}} = \left\langle \mathbf{K}_{\mathbf{g}|\mathbf{f}} \right\rangle_{\mathbf{f}} = \text{diag}\,(\overline{\overline{\mathbf{g}}})\,,$$

Now look at second term. Recall that the object is a random process $f(\mathbf{r})$ and hence described by an autocovariance function:

$$k_{\mathbf{f}}(\mathbf{r}, \mathbf{r}') = \left\langle [f(\mathbf{r}) - \overline{f}(\mathbf{r})]\,[f(\mathbf{r}') - \overline{f}(\mathbf{r}')]\,\right\rangle\ .$$

The autocovariance function can be regarded as the kernel of an integral operator $\mathcal{K}_{\mathbf{f}}$, and the second term in the decomposition can be written formally as

$$\mathbf{K}_{\overline{\mathbf{g}}}^{obj} = \boldsymbol{\mathcal{H}}\mathcal{K}_{\mathbf{f}}\boldsymbol{\mathcal{H}}^{\dagger}\,.$$

The covariance decomposition in practice

$$\mathbf{K_g} = \overline{\mathbf{K}}_{\mathbf{g}}^{Pois} + \mathbf{K}_{\overline{\mathbf{g}}}^{obj} = \text{diag}\,(\overline{\overline{\mathbf{g}}}) + \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{K}}_{\mathbf{f}}\boldsymbol{\mathcal{H}}^{\dagger}\,.$$

This all looks very formidable, but in practice both terms can be estimated by Monte Carlo methods:

1. Generate lots of noise-free simulations of objects on a pixel grid

2. For each simulated object, generate corresponding noise-free image

3. Sample mean image gives estimate of $\overline{\mathbf{K}}_{\mathbf{g}}^{Pois}$

4. Sample covariance matrix is estimate of $\mathbf{K}_{\overline{\mathbf{g}}}^{obj}$

5. Resulting sum of two terms is full rank and hence invertible!

We do this routinely in the Center for Gamma-ray imaging

Application to adaptive optics

Game plan:

Enumerate all sources of randomness

Carefully state all statistical assumptions

Account for closed loop

Develop overall spatiotemporal covariance matrix for image data

Use the model to discuss task performance

Components of an AO system: Their inputs and outputs

- Wavefront sensor (nonlinear CD)
  Input: Wavefront for k^{th} frame: $\phi^{(k)}(\mathbf{r})$, $\mathbf{r} = (x, y)$
  Output: Set of sensor signals, $\{v_m^{(k)}\} \equiv \{\mathbf{v}^{(k)}\} \equiv \mathbf{V}$.
- Estimator (not really a wavefront reconstructor) (Nonlinear, DD)
  Input: Vector of sensor signals for one frame, $\mathbf{v}^{(k)}$
  (Could include a predictor, then would use preceding frames also)
  Output: Estimates of wavefront parameters for that frame: $\hat{\boldsymbol{\alpha}}^{(k)}$.
- Control (Usually linear, DD)
  Input: Estimates of wavefront parameters: $\hat{\boldsymbol{\alpha}}^{(k)}, \hat{\boldsymbol{\alpha}}^{(k-1)}, \hat{\boldsymbol{\alpha}}^{(k-2)}, \ldots$
  Output: Signals to deformable mirror on next frame: $\boldsymbol{\beta}^{(k+1)}$
- Deformable mirror (Linear, DC/CC)
  Inputs: Wavefront and control signals
  Output: Corrected wavefront (complex pupil function $t^{(k)}(\mathbf{r})$)
- Science camera (Linear, CD)
  Inputs: Sky radiance, incoherent PSF derived from $t^{(k)}(\mathbf{r})$.
  Output: Set of pixel outputs for each frame: $\{g_m^{(k)}\} \equiv \{\mathbf{g}^{(k)}\} \equiv \mathbf{G}$.

Triply stochastic analysis

Consider the main path:

Object – atmosphere – deformable mirror – science camera

Sources of randomness:

Object $\mathbf{f}$

Incoherent PSF $\mathbf{p}$ (square mod of FT of random complex pupil function)

Detector noise $\mathbf{n}$ (photon + readout)

These three factors contribute to the statistics of the science image $\mathbf{g}$

Role of the WFS and feedback loop: Control the incoherent PSF.

Detail: Relation of PSF to wavefront

Start with well-known relation between coherent PSF and complex pupil function:

$$p_{coh}(\mathbf{r}) = \text{const} \int_{ap} d^2r' \, t_{pup}(\mathbf{r}') \, \text{exp} \left[-2\pi i \frac{\mathbf{r} \cdot \mathbf{r}'}{\lambda f} \right] ,$$

where f is focal length, object is at ∞ and PSF is measured in focal plane.

Model the atmosphere and DM as thin phase plates in pupil:

$$t_{pup}(\mathbf{r}) = \text{exp} \left\{ i \left[\phi_{atm}(\mathbf{r}) - \phi_{dm}(\mathbf{r}) \right] \right\} .$$

Incoherent PSF given by

$$p(\mathbf{r}) = \text{const} \left| p_{coh}(\mathbf{r}) \right|^2$$

PSF $p(\mathbf{r})$ is random since $\phi_{atm}(\mathbf{r})$ and $\phi_{dm}(\mathbf{r})$ are random. Can express $\phi_{atm}(\mathbf{r})$ in terms of random Zernike coefficients and $\phi_{dm}(\mathbf{r})$ in terms of random coefficients for mirror influence functions.

Triply stochastic averaging for a single frame

Let $\mathbf{g}$ be a single frame of data from the science camera, with statistics that depend on the incoherent PSF $\mathbf{p}$ and noise realization $\mathbf{n}$ for that frame.

Overall average

$$\overline{\overline{\overline{\mathbf{g}}}} = \langle\langle\langle\, \mathbf{g}\,\rangle_{\mathbf{g}|\mathbf{p},\mathbf{f}}\rangle_{\mathbf{p}|\mathbf{f}}\rangle_{\mathbf{f}}\ .$$

Comments:

1. Inner average is over just detector noise. Easy since noise for different pixels is statistically independent and $\text{pr}(\mathbf{g}|\mathbf{p},\mathbf{f}\,) = \text{pr}[\mathbf{g}|\overline{\mathbf{g}}(\mathbf{p},\mathbf{f}\,)]$.

2. Second average involves $\text{pr}(\mathbf{p}|\mathbf{f}\,)$. Requires detailed knowledge or simulation of WFS and control loop.

3. Third average involves $\text{pr}(\mathbf{f}\,)$. Requires model or simulation code for sky background .

A potential simplification

The second average, $\langle \cdots \rangle_{\mathbf{p}|\mathbf{f}}$ would be a lot easier if we could assume that the incoherent PSF $\mathbf{p}$ was independent of the object $\mathbf{f}$ being imaged.

Independence assumption might be OK if there is a strong guide star well separated from the science object.

$$\mathbf{f} = \mathbf{f}_{sci} + \mathbf{f}_{bg} + \mathbf{f}_{gs}$$

Might assume that PSF depends only on the guide star, $\mathbf{f}_{gs}$.

Independence would not hold if we were looking for a planetary companion to the guide star itself. (But why would we want to flatten the wavefront in that case?)

Independence assumption would be terrible for solar imaging

Slight digression: From PSF to sensitivity function

We don't really want to talk about a PSF for a digital imaging system (too conventional).

Instead we use a sensitivity function $h_m(\mathbf{r})$, defined so that

$$\overline{g}_m = \int d^2r\, h_m(\mathbf{r})\, f(\mathbf{r})\,.$$

If $I(\mathbf{r})$ is the irradiance in the detector plane (for some particular object and PSF), we can also write

$$\overline{g}_m = \int d^2r'\, w_m(\mathbf{r}')\, I(\mathbf{r}')\,,$$

where $w_m(\mathbf{r})$ is aperture of m^{th} detector. For shift-invariant CC imaging

$$\overline{g}_m = \int d^2r'\, w_m(\mathbf{r}') \int d^2r\, p(\mathbf{r}' - \mathbf{r})\, f(\mathbf{r})$$

so

$$h_m(\mathbf{r}) = \int d^2r'\, w_m(\mathbf{r}')\, p(\mathbf{r}' - \mathbf{r})\,.$$

Key point: Sensitivity function is now random since PSF is.

$\boxed{\text{Partial means}}$

If we average over detector noise alone, then

$$\overline{g}_m = \overline{g}_m(\mathbf{p}, \mathbf{f}) = \int d^2r \, h_m(\mathbf{r}) \, f(\mathbf{r}) \, .$$

Next average is over random PSF $\mathbf{p}$ given $\mathbf{f}$:

$$\overline{\overline{g}}_m = \overline{g}_m(\mathbf{f}) = \int d^2r \, \overline{h}_m(\mathbf{r}) \, f(\mathbf{r}) \, .$$

Final average is over object variability. If $\mathbf{p}$ is independent of $\mathbf{f}$, then

$$\overline{\overline{\overline{g}}}_m = \int d^2r \, \overline{h}_m(\mathbf{r}) \, \overline{f}(\mathbf{r}) \, .$$

If independence does not hold, must write

$$\overline{\overline{\overline{g}}}_m = \int d^2r \, \left\langle \overline{h}_m(\mathbf{r}, \mathbf{f}) \, f(\mathbf{r}) \right\rangle_{\mathbf{f}} \, .$$

Triply stochastic decomposition of the covariance matrix

Definition of overall covariance:

$$\mathbf{K_g} \equiv \langle [\mathbf{g} - \overline{\overline{\overline{\mathbf{g}}}}][\mathbf{g} - \overline{\overline{\overline{\mathbf{g}}}}]^t \rangle_{\mathbf{g},\mathbf{p},\mathbf{f}} = \langle\langle\langle [\mathbf{g} - \overline{\overline{\overline{\mathbf{g}}}}][\mathbf{g} - \overline{\overline{\overline{\mathbf{g}}}}]^t \rangle_{\mathbf{g}|\mathbf{p},\mathbf{f}} \rangle_{\mathbf{p}|\mathbf{f}} \rangle_{\mathbf{f}} \,.$$

Now add and subtract terms in each factor:

$$\mathbf{K_g} = \langle\langle\langle [\mathbf{g} - \overline{\mathbf{g}} + \overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}} + \overline{\overline{\mathbf{g}}} - \overline{\overline{\overline{\mathbf{g}}}}]\, [\mathbf{g} - \overline{\mathbf{g}} + \overline{\mathbf{g}} - \overline{\overline{\mathbf{g}}} + \overline{\overline{\mathbf{g}}} - \overline{\overline{\overline{\mathbf{g}}}}]^t \rangle_{\mathbf{g}|\mathbf{p},\mathbf{f}} \rangle_{\mathbf{p}|\mathbf{f}} \rangle_{\mathbf{f}} \,.$$

All cross terms vanish (without independence assumptions) and we get

$$\mathbf{K_g} = \overline{\overline{\mathbf{K}}}_{\mathbf{g}}^{noise} + \overline{\mathbf{K}}_{\overline{\mathbf{g}}}^{PSF} + \mathbf{K}_{\overline{\overline{\mathbf{g}}}}^{obj}$$

$\overline{\overline{\mathbf{K}}}_{\mathbf{g}}^{noise}$: Diagonal matrix from readout and Poisson noise, averaged over $\mathbf{p}$ and $\mathbf{f}$

$\overline{\mathbf{K}}_{\overline{\mathbf{g}}}^{PSF}$: Contribution from random PSF, averaged over objects unless $\mathbf{p}$ is independent of $\mathbf{f}$

$\mathbf{K}_{\overline{\overline{\mathbf{g}}}}^{obj}$: Contribution from object randomness

Scale of the correlations

We now have three separate covariance components. What is the range of the correlations for each?

Noise term: Zero range (delta-correlated)

PSF term: Correlation range $\simeq$ width of *uncorrected* PSF

Object term: Can be long-range or multi-scale, often modeled as fractal

$\boxed{\text{Howdaya get the component covariances?}}$

Basic answer: Monte-Carlo simulation of noise-free images of random objects with random PSFs

Form overall mean of these images and estimate noise covariance by

$$\overline{\overline{\mathbf{K}}}_{\mathbf{g}}^{noise} = \text{diag}\left[\overline{\overline{\overline{\mathbf{g}}}} + \sigma_{read}^{2}\right]$$

Other two terms handled similarly:

$\overline{\mathbf{K}}_{\overline{\mathbf{g}}}^{PSF}$: Estimate from sample covariance for sensitivity function plus mean objects.

$\mathbf{K}_{\overline{\overline{\mathbf{g}}}}^{obj}$: Estimate from sample covariance on object plus mean sensitivity functions

Details on PSF and object terms

For reference:

$$\left[\overline{\mathbf{K}}_{\overline{\mathbf{g}}}^{PSF}\right]_{mm'} = \int d^2r \int d^2r' \left\langle f(\mathbf{r})\, f(\mathbf{r}') \right\rangle \left\langle \Delta h_m(\mathbf{r})\, \Delta h_{m'}(\mathbf{r}') \right\rangle$$

$$\left[\mathbf{K}_{\overline{\overline{\mathbf{g}}}}^{obj}\right]_{mm'} = \int d^2r \int d^2r' \left\langle \Delta f(\mathbf{r})\, \Delta f(\mathbf{r}') \right\rangle \left\langle h_m(\mathbf{r})\, h_{m'}(\mathbf{r}') \right\rangle \ ,$$

where $\Delta \Rightarrow$ "subtract off the mean"

Computational simplification: $h_m(\mathbf{r})$ and $\Delta h_m(\mathbf{r})$ will be sharply peaked near $\mathbf{r} = \mathbf{r}_m$.

Statistics of the control loop

- Basic trick: assume loop is closed and phase excursions are small

- Relate covariance of incoherent PSF to covariance of net phase by Taylor expansion

- Determine covariance of net phase from atmospheric covariance (Zernikes), covariance of output of WF estimator, and control algorithm. (Need to solve difference equation or treat problem as Markov chain.)

- Use generalized CR bound for covariance of output of WF estimator

$$\mathbf{K}_{\hat{\boldsymbol{\alpha}}} \geq \mathbf{F}^{-1}$$

where $\mathbf{F}$ is Fisher information matrix (Lecture 10) for estimation of wavefront parameters $\boldsymbol{\alpha}$. Equality holds for efficient estimator, holds asymptotically for MLE.

$\boxed{\text{Still to do}}$

- Consider computational issues in estimate of covariance components

- Fill in the details for statistics of the control loop

- Work out details of case where PSF depends on the science object.

- Extend theory to spatiotemporal covariances

- Pick a task and work out FOM

In lieu of conclusions: Questions for the group

What do you think about the assumption that the random PSF is independent of the science object, at least for stellar AO?

Can Thomas's simulation code be used to generate sample PSFs and hence sample sensitivity functions? Can it generate sample covariances?

Can Ruth's ferroelectric light valve be used similarly?

Does it make sense to do the initial analysis with no memory in the control loop?

What tasks are most important in astronomical AO?

What are the important system parameters?

Should we consider the science object and the sky background to be time-independent?