# Introduction to Multivariate Speckle Statistics

## The statistics in statistical decision theory: What do you need to know and how do you get it?

Harrison H. Barrett
College of Optical Sciences
and Dept. of Radiology
University of Arizona

# Outline

- Key points from lecture 1
  - Scalar test statistics
  - ROC and LROC CURVES
- Optimum linear (Hotelling) test statistic for image sequences
- Statistical models of AO systems
- Analytical and simulation approaches to multivariate speckle statistics
- Summary and conclusions
- Revised proposal for the Exoplanet Challenge

# Notation and simplifications

- Given a data set $\mathbf{G}$ (sequence of $J$ images)

  $$\mathbf{G} = \{\, \mathbf{g}^{(j)},\ j = 1,\ldots,\ J \,\}$$

- Two possible hypotheses about the object:
  - Signal absent, $S_-$
  - Signal present, $S_+$
- Make a statistically-based determination of whether the signal is present or absent, then determine its location
- To simplify today's discussion, consider only:
  - Hotelling observer and ROC curve
  - Scanning Hotelling observer and LROC curve
- Sources of randomness considered today
  - Detector readout noise and Photon (Poisson) noise
  - Random system PSF

Assume that a binary decision (signal-present or signal-absent) must be made for every image and that there is no randomness in the decision (Don't guess; don't equivocate)

With these assumptions, the decision strategy is:

To decide if a signal is present at a particular location $\mathbf{r}_0$, use

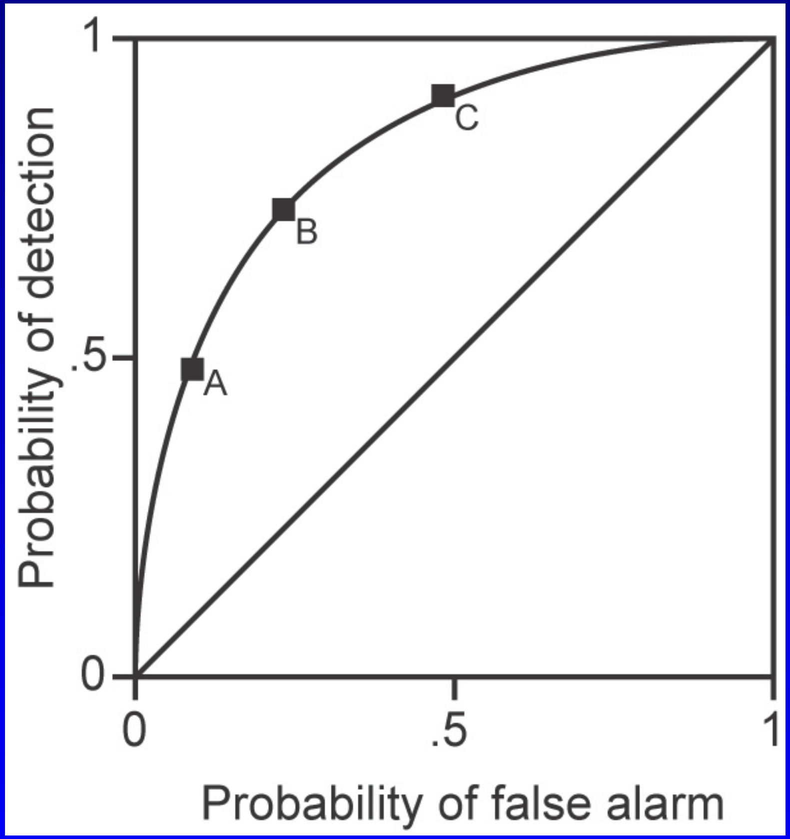
$$t(\mathbf{G}; \mathbf{r}_0) \underset{D_-}{\overset{D_+}{\gtrless}} \text{threshold}$$

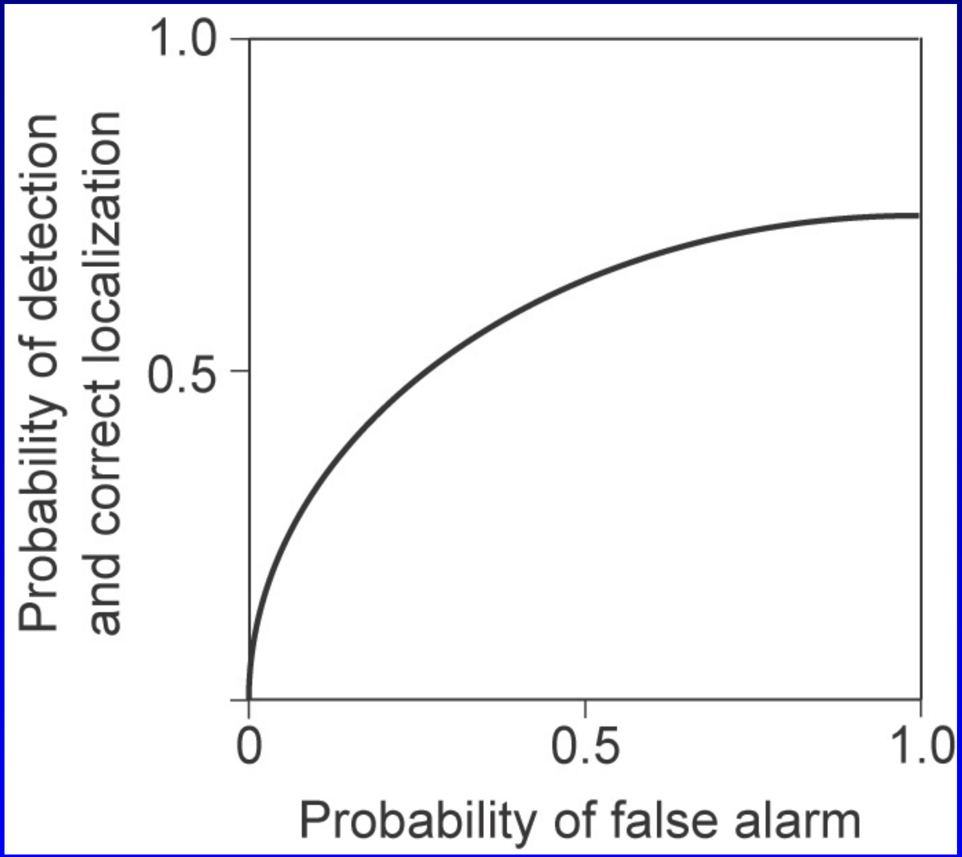
This inequality is to be read:
"decide signal present when the greater-than sign holds;
decide hypothesis signal absent when the less-than sign holds."

To decide if a signal is present anywhere in the data, use

$$\max_{\mathbf{r}_0} \; t(\mathbf{G}; \mathbf{r}_0) \underset{D_-}{\overset{D_+}{\gtrless}} \text{threshold}$$

From a set of trial images with known truth, can construct ROC and LROC curves for any test statistic $t(\mathbf{G}; \mathbf{r}_0)$ , (i.e. any detection algorithm)




Receiver Operating Characteristic

Points A, B and C correspond
to different decision thresholds

Localization ROC

Must specify tolerance on
required localization accuracy

# The Hotelling observer

- Performs only linear operations on data
- Maximizes SNR for test statistic
- Requires knowledge of ensemble means and covariance matrices of the images
- Equivalent to ideal observer for Gaussian data

3080 J. Opt. Soc. Am. A/Vol. 23, No. 12/December 2006

# Objective assessment of image quality. IV. Application to adaptive optics

This paper presents the full derivation of the mean and covariance of the data set **G** for object and system PSF being general spatiotemporal random processes

Here we summarize the results for a simple object (a star) which is not random and does not vary with time and where the PSF randomness arises from a phase perturbation in the pupil

## Two cases

- Simple AO system, pinned speckle
- Lyot coronagraph

# Notation for object and image

Object model (host star at origin, possible planet at $\mathbf{r}_0$):

Signal absent: $f(\mathbf{r}) = A\delta(\mathbf{r})$

Signal present: $f(\mathbf{r}) = A\delta(\mathbf{r}) + a\delta(\mathbf{r} - \mathbf{r}_0)$

General linear mapping of object to discrete data (value in image frame $j$ at pixel $m$):

$$g_m^{(j)} = \int d^2r\, f(\mathbf{r}) h_m^{(j)}(\mathbf{r}) + n_m^{(j)},$$

where $h_m^{(j)}(\mathbf{r})$ is integral of PSF over frame $j$ and pixel $m$ and $n_m^{(j)}$ is noise

Signal absent: $g_m^{(j)} = A h_m^{(j)}(0) + n_m^{(j)}$

Signal present: $g_m^{(j)} = A h_m^{(j)}(0) + a h_m^{(j)}(\mathbf{r}_0) + n_m^{(j)}$

Shorthand: $\mathbf{G} = \boldsymbol{\mathcal{H}}\mathbf{f} + \mathbf{n}$

# Doubly stochastic averaging: mean data

General spatiotemporal data: $\mathbf{G} = \boldsymbol{\mathcal{H}}\mathbf{f} + \mathbf{n}$

Average over zero-mean noise: $\overline{\mathbf{G}} \equiv \langle \mathbf{G} \rangle_{\mathbf{G}|\mathcal{H}} = \boldsymbol{\mathcal{H}}\mathbf{f}$

Average over random PSFs: $\overline{\overline{\mathbf{G}}} \equiv \langle \langle \mathbf{G} \rangle_{\mathbf{G}|\mathcal{H}} \rangle_{\mathcal{H}} = \overline{\boldsymbol{\mathcal{H}}}\mathbf{f}$

Key point: these averages are ensemble averages,
over an infinite set of realizations of noise and PSFs,
not time averages or sample averages over a few realizations

Overall covariance matrix (details on next slide)

$$\mathbf{K_G} \equiv \left\langle \left\langle \left(\mathbf{G} - \overline{\overline{\mathbf{G}}}\right) \left(\mathbf{G} - \overline{\overline{\mathbf{G}}}\right)^{\dagger} \right\rangle_{\mathbf{G}|\mathcal{H}} \right\rangle_{\mathcal{H}} = \overline{\mathbf{K}}_{\mathbf{G}}^{noise} + \mathbf{K}_{\overline{\mathbf{G}}}^{PSF}$$

Overall covariance is sum of two terms, with
no assumption that noise and PSF are statistically independent

Noise term includes Poisson noise and readout noise from
science camera; it is almost always a diagonal matrix

Random PSF induces correlations from frame to frame and from pixel to pixel,
and understanding these correlations is key to optimal planet detection

Useful assumption for exoplanet detection:
$\mathbf{K_G}$ is the same for signal-present and signal absent
because the planet signal is so weak

# Some algebra, which you can read at your leisure

In outer-product notation, the overall covariance matrix is given by

$$\mathbf{K}_{\mathbf{G}} \equiv \left\langle \left\langle \left(\mathbf{G} - \overline{\overline{\mathbf{G}}}\right) \left(\mathbf{G} - \overline{\overline{\mathbf{G}}}\right)^{\dagger} \right\rangle_{\mathbf{G}|\mathcal{H}} \right\rangle_{\mathcal{H}}$$

where the dagger denotes adjoint (transpose for a simple real matrix). In component form,

$$[\mathbf{K}_{\mathbf{G}}]_{m,m'}^{j,j'} \equiv \left\langle \left\langle \left(G_m^{(j)} - \overline{\overline{G}}_m^{(j)}\right) \left(G_{m'}^{(j')} - \overline{\overline{G}}_{m'}^{(j')}\right) \right\rangle_{\mathbf{G}|\mathcal{H}} \right\rangle_{\mathcal{H}}$$

Now add and subtract $\overline{\mathbf{G}}$ in first equation:

$$\mathbf{K}_{\mathbf{G}} \equiv \left\langle \left\langle \left(\mathbf{G} - \overline{\mathbf{G}} + \overline{\mathbf{G}} - \overline{\overline{\mathbf{G}}}\right) \left(\mathbf{G} - \overline{\mathbf{G}} + \overline{\mathbf{G}} - \overline{\overline{\mathbf{G}}}\right)^{\dagger} \right\rangle_{\mathbf{G}|\mathcal{H}} \right\rangle_{\mathcal{H}}$$

$$= \left\langle \left\langle (\mathbf{G} - \overline{\mathbf{G}}) (\mathbf{G} - \overline{\mathbf{G}})^{\dagger} \right\rangle_{\mathbf{G}|\mathcal{H}} \right\rangle_{\mathcal{H}} + \left\langle \left(\overline{\mathbf{G}} - \overline{\overline{\mathbf{G}}}\right) \left(\overline{\mathbf{G}} - \overline{\overline{\mathbf{G}}}\right)^{\dagger} \right\rangle_{\mathcal{H}} \equiv \overline{\mathbf{K}}_{\mathbf{G}}^{noise} + \mathbf{K}_{\overline{\mathbf{G}}}^{PSF}$$

Note that the cross term has vanished identically since

$$\left\langle \left\langle (\mathbf{G} - \overline{\mathbf{G}}) \left(\overline{\mathbf{G}} - \overline{\overline{\mathbf{G}}}\right)^{\dagger} \right\rangle_{\mathbf{G}|\mathcal{H}} \right\rangle_{\mathcal{H}} = \left\langle \left\langle (\mathbf{G} - \overline{\mathbf{G}}) \right\rangle_{\mathbf{G}|\mathcal{H}} \left(\overline{\mathbf{G}} - \overline{\overline{\mathbf{G}}}\right)^{\dagger} \right\rangle_{\mathcal{H}} = 0$$

# Form of the Hotelling observer for detection of a planet at $\mathbf{r}_0$

$$t(\mathbf{G}) = \Delta\overline{\overline{\mathbf{G}}}^{\dagger} \left[ \overline{\mathbf{K}}_{\mathbf{G}}^{noise} + \mathbf{K}_{\overline{\mathbf{G}}}^{PSF} \right]^{-1} \mathbf{G}$$

$$\left[ \Delta\overline{\overline{\mathbf{G}}} \right]_{m}^{(j)} = a\overline{h}_{m}^{(j)}(\mathbf{r}_0)$$

- Preprocess data sequence with inverse of overall covariance
- Take scalar product with mean planet signal
- Compare resulting test statistic to a threshold
  or scan to find most probable planet location

# Modeling the PSF (speckle) term in the covariance matrix

- Monte Carlo methods
  - Simulate sample functions of $\phi(\mathbf{r}, t)$ (potentially non-Gaussian)
  - Propagate to image plane, compute irradiance $I(\mathbf{r}, t)$
  - Compute sample covariance matrix

- Analytical approach
  - Treat pupil phase $\phi(\mathbf{r}, t)$ as spatiotemporal Gaussian random process
  - Assume weak phase, use pinned-speckle theory
  - Propagate resulting complex field to image plane
  - Compute characteristic functional of irradiance
  - Derive covariance matrix from characteristic functional
  - Derive any other desired statistical property from characteristic functional

For a short introduction to characteristic functionals, see
http://media.nakfi.org/2010/tutorials/harrison_barrett/harrison_barrett.html

January 15, 2004 / Vol. 29, No. 2 / OPTICS LETTERS 159

# Anomalous intensity of pinned speckles at high adaptive correction

E. E. Bloemhof

THE ASTROPHYSICAL JOURNAL, 596:702–712, 2003 October 10


# THE STRUCTURE OF HIGH STREHL RATIO POINT-SPREAD FUNCTIONS

MARSHALL D. PERRIN[1]
Astronomy Department, University of California, Berkeley, CA 94720

ANAND SIVARAMAKRISHNAN AND RUSSELL B. MAKIDON
Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218

BEN R. OPPENHEIMER
Astrophysics Department, American Museum of Natural History,
Central Park West at 79th Street, New York, NY 10024

AND

JAMES R. GRAHAM
Astronomy Department, University of California, Berkeley, CA 94720
*Received 2003 March 28; accepted 2003 June 20*

Pupil function with random phase aberration

$$u_{pup}(\mathbf{r},t) = a_{ap}(\mathbf{r}) \exp[i\phi(\mathbf{r},t)],$$

where $\mathbf{r}$ is a 2D vector in the pupil plane and $a_{ap}(\mathbf{r})$ is the transmittance of the pupil aperture.

The random field incident in the back focal plane is given in the Fresnel approximation by

$$u(\mathbf{r},t) = \frac{1}{i\lambda f} \exp\left(\frac{i\pi}{\lambda f} r_d^2\right) \int_{ap} d^2r' \exp[i\phi(\mathbf{r}',t)] \exp\left(-\frac{2\pi i}{\lambda f}\mathbf{r}'\cdot\mathbf{r}\right),$$

Expand the focal-plane field as

$$u(\mathbf{r},t) = \frac{1}{i\lambda f} \exp\left(\frac{i\pi}{\lambda f} r_d^2\right) \int_{ap} d^2r' \left[1 + i\phi(\mathbf{r}',t) - \tfrac{1}{2}\phi^2(\mathbf{r}',t) + \cdots\right] \exp\left(-\frac{2\pi i}{\lambda f}\mathbf{r}'\cdot\mathbf{r}\right).$$

Retaining terms only up to second order in $\phi(\mathbf{r}',t)$, we can write the random focal-plane irradiance as

$$I(\mathbf{r},t) \approx \frac{1}{\lambda^2 f^2}\left[A_{ap}\left(\frac{\mathbf{r}}{\lambda f}\right)\right]^2 + \frac{2}{\lambda^2 f^2} A_{ap}\left(\frac{\mathbf{r}_d}{\lambda f}\right) \int_{ap} d^2r' \phi(\mathbf{r}',t) \sin\left(\frac{2\pi}{\lambda f}\mathbf{r}'\cdot\mathbf{r}\right)$$

$$+\frac{1}{\lambda^2 f^2}\int_{ap} d^2r' \int_{ap} d^2r'' \phi(\mathbf{r}',t)\phi(\mathbf{r}'',t) \exp\left(\frac{2\pi i}{\lambda f}(\mathbf{r}'-\mathbf{r}'')\cdot\mathbf{r}\right) - \frac{1}{\lambda^2 f^2} A_{ap}\left(\frac{\mathbf{r}_d}{\lambda f}\right) \int_{ap} d^2r' \phi^2(\mathbf{r}',t) \cos\left(\frac{2\pi}{\lambda f}\mathbf{r}'\cdot\mathbf{r}\right).$$

In terms of spatial Fourier transforms:

$$I(\mathbf{r},t) = \frac{1}{\lambda^2 f^2}\Big\{ |A_{ap}(\tilde{\mathbf{r}})|^2 - 2A_{ap}(\tilde{\mathbf{r}})\mathrm{Im}\left\{[A_{ap} * \Phi](\tilde{\mathbf{r}})\right\} + |[A_{ap} * \Phi](\tilde{\mathbf{r}})|^2 - A_{ap}(\tilde{\mathbf{r}})\mathrm{Re}\left\{[A_{ap} * \Phi * \Phi](\tilde{\mathbf{r}})\right\} \Big\},$$

where $\Phi(\boldsymbol{\rho},t)$ is the 2D Fourier transform of $\phi(\mathbf{r}',t)$ and $\tilde{\mathbf{r}} \equiv \mathbf{r}/(\lambda f)$.

Equations of pinned speckle on last slide are just a start. For the Hotelling observer we need the ensemble means and covariance matrices of the random focal-plane irradiance:

$$\overline{I}(\mathbf{r},t) = \langle I(\mathbf{r},t)\rangle \ , \qquad K_I(\mathbf{r},t;\mathbf{r}',t') = \left\langle \left[I(\mathbf{r},t) - \overline{I}(\mathbf{r},t)\right] \left[I(\mathbf{r}',t') - \overline{I}(\mathbf{r}',t')\right] \right\rangle \ ,$$

where $\langle \cdots \rangle$ denotes ensemble average

To evaluate these averages, we need a model for the multivariate statistics of the pupil phase. Possibilities include:

- Treat $\phi(\mathbf{r}, t)$ as a real Gaussian random process
- Perhaps add assumption of spatial stationarity
- Perhaps model time dependence as Taylor frozen flow

## Rationale for Gaussian phase models

- Kolmogorov model is Gaussian, so residual phase after AO still Gaussian
- Drive signals to DM can be noisy, but Gaussian model should be good
- Gaussian random processes are fully determined by their mean and covariance functions
- Can find characteristic functional for focal-plane irradiance if pupil phase is a Gaussian random process
- Spatial stationarity of the residual pupil phase might work for high-order AO and atmospheric phase perturbations

## Potential drawbacks to Gaussian phase assumption

- Statistics of quasistatic phase perturbations unknown, may not be Gaussian
- Quasistatic phase definitely not stationary

# Example of a speckle covariance calculated from a characteristic functional

Stationary Gaussian model for pupil phase, no time variable

$S_\phi(\boldsymbol{\rho})$ = power spectral density of phase

$$K_\mathbf{I}(\mathbf{r}_1,\mathbf{r}_2) = -\frac{2}{\lambda^4 f^4} A_{ap}\left(\frac{\mathbf{r}_1}{\lambda f}\right) A_{ap}\left(\frac{\mathbf{r}_2}{\lambda f}\right)\left[S_\phi\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2\lambda f}\right) A_{ap}\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{\lambda f}\right) - S_\phi\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{2\lambda f}\right) A_{ap}\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{\lambda f}\right)\right]$$

$$-4\left(\frac{A_{ap}(0)}{\lambda^2 f^2}\right)^2 S_\phi\left(\frac{\mathbf{r}_1}{\lambda f}\right) S_\phi\left(\frac{\mathbf{r}_2}{\lambda f}\right)\left[A_{ap}\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{\lambda f}\right)\right]^2$$

$$-\frac{4}{\lambda^4 f^4} A_{ap}\left(\frac{\mathbf{r}_1}{\lambda f}\right) A_{ap}\left(\frac{\mathbf{r}_2}{\lambda f}\right)\left\{\left[A_{ap}\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{\lambda f}\right)+A_{ap}\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{\lambda f}\right)\right]\left[[S_\phi \star S_\phi]\left(\frac{\mathbf{r}_1-\mathbf{r}_2}{\lambda f}\right)+[S_\phi \star S_\phi]\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{\lambda f}\right)\right]\right\}$$

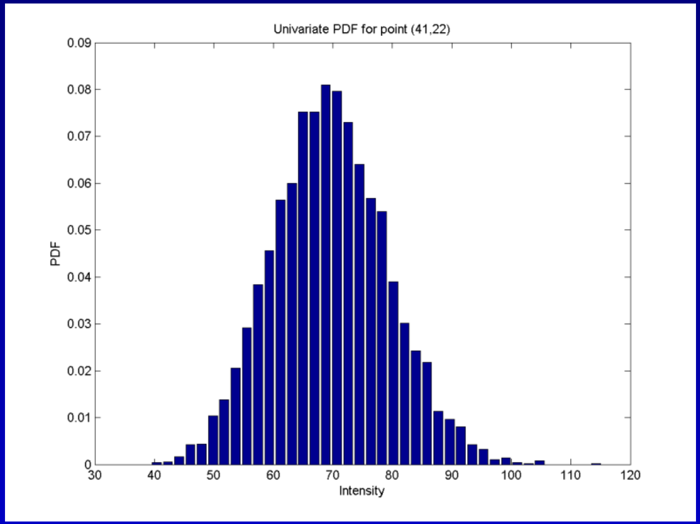
$$+\frac{4}{\lambda^4 f^4}\left[A_{ap}\left(\frac{\mathbf{r}_1}{\lambda f}\right)\right]^2\left[S_\phi\left(\frac{\mathbf{r}_2}{\lambda f}\right)\right]^2 + \frac{4}{\lambda^4 f^4}\left[A_{ap}\left(\frac{\mathbf{r}_2}{\lambda f}\right)\right]^2\left[S_\phi\left(\frac{\mathbf{r}_1}{\lambda f}\right)\right]^2$$

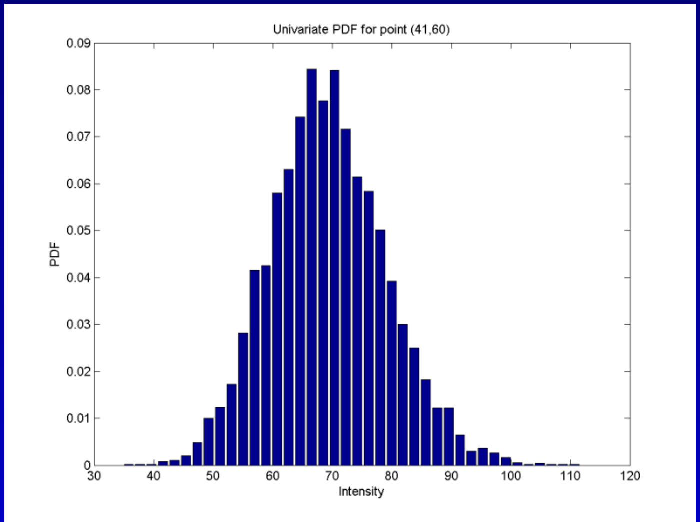
For AO system with very high Strehl ratio, can neglect lines 2 - 4

Math then shows that $K_\mathrm{I}(\mathbf{r}_1, \mathbf{r}_1) = -K_\mathrm{I}(\mathbf{r}_1, -\mathbf{r}_1)$

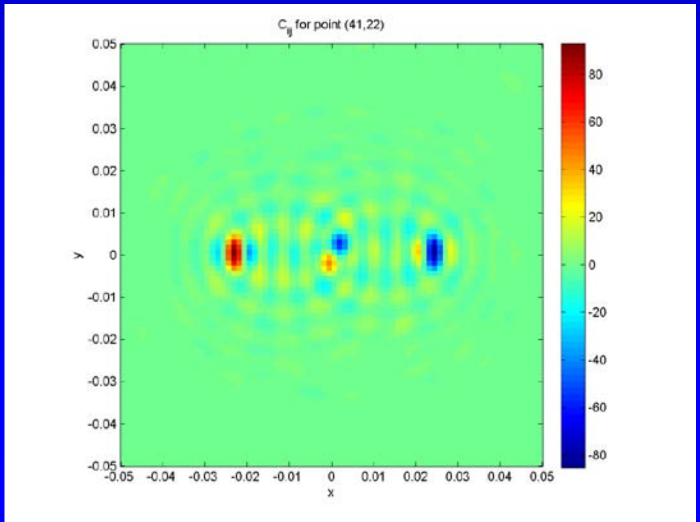
In this same limit, focal-plane irradiance is a Gaussian random process and the Hotelling observer is ideal

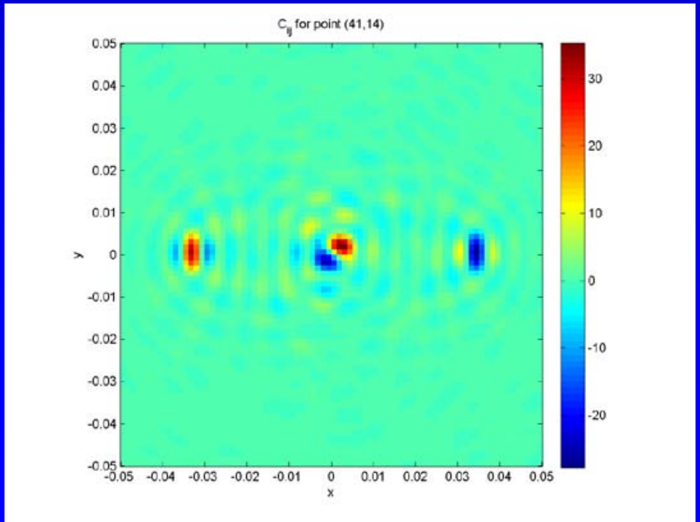
# Monte Carlo simulation for same speckle model as last slide (Strehl ratio = 0.91)
Work of Julia Sakamoto




Univariate PDFs for irradiance at two different points in the focal plane




Covariance maps, $K_I(x_0, y_0; x,y)$ for two fixed locations $(x_0, y_0)$

# Fast computation of Lyot-style coronagraph propagation

R. Soummer*[1,2], L. Pueyo[3],
A. Sivaramakrishnan [1,2,4], and R. J. Vanderbei[5]

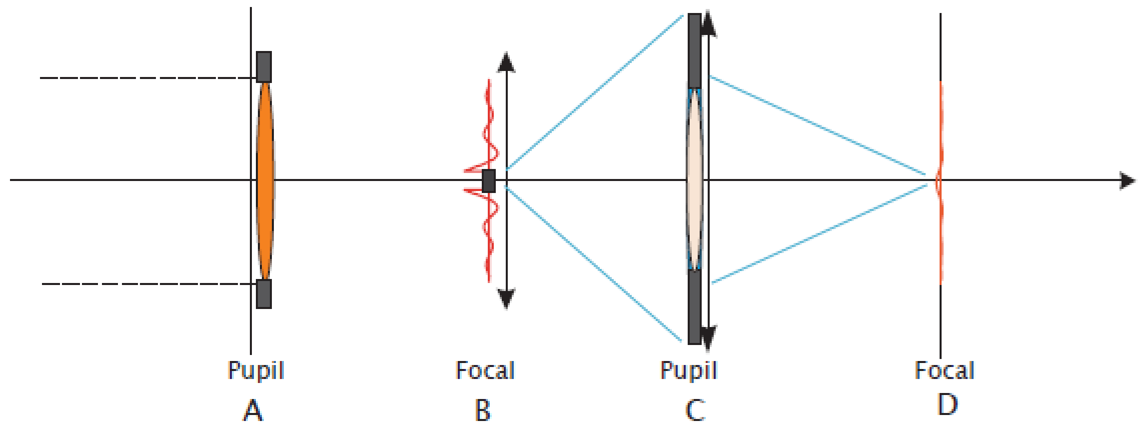



Fig. 1. Illustration of the four coronagraphic planes: the pupil corresponds to Plane A (possibly apodized). A focal masks (hard-edged, or phase mask) is placed in the focal plane B, and a Lyot stop (possibly undersized) in plane C.

Can readily extend calculations from previous slides to this configuration;
just replace Fourier transform with coherent propagator from plane A to plane D.

# Summary

- Statistical approaches to evaluation of exoplanet-detection algorithms must be based on average probability of detection and false-alarm rate over many images
- ROC and LROC curves are a well-established way of displaying the results
- Ideal observers for ROC and LROC are known, require multivariate PDFs for huge data vectors
- Ideal linear (Hotelling) observers also known, require means and covariance matrices of the image data
- Data are spatiotemporal (image sequences), so covariances must include space and time variables
- Both analytical and simulation methods exist for finding the requisite multivariate statistics
- Both theory and simulation show that focal-plane irradiance is Gaussian if Strehl ratio is large; Hotelling is ideal in this case
- Once the means and covariances are either known or sampled, Hotelling and scanning-Hotelling observers can be constructed and evaluated (Luca Caucci)

# Revised proposal for the workshop challenge

- Make it a two-stage challenge
- Stage 1 (pre-workshop)
  - Simple telescope model, DM conjugate to pupil
  - Can include Lyot coronagraph
  - Can include quasistatic speckle as well as residual atmospheric speckle, but model it as originating in the pupil
  - Use simulated data (need ~ 100 data sets, each an image sequence)
- Stage 2 (post-workshop)
  - State-of-the-art telescope
  - Include current coronagraph, electric-field conjugation, apodization,...
  - Consider use of real data with added planets
- In both stages, perform detection, localization and magnitude estimation; analyze results by ROC and LROC

Supplemental slides on characteristic functionals
in case further discussion is needed

http://media.nakfi.org/2010/tutorials/harrison_barrett/harrison_barrett.html

# Characteristic functionals

An object $\mathbf{f}$ is a function ...

... hence a vector in an infinite-dimensional space

Thus its PDF $pr(\mathbf{f})$ is infinite-dimensional ...

which is difficult to conceive, much less write down.

But an infinite-dimensional characteristic *functional* always exists, and can often be expressed *in a simple analytic form*

This characteristic functional contains *all possible statistical information* about the object function

# Characteristic functional -- definition

Recall the definition of the characteristic *function* for a $M$D real random vector:

$$\psi_{\mathrm{g}}(\xi) = \left\langle \exp(-2\pi i \xi^t \mathrm{g}) \right\rangle , \qquad (8.26)$$

Here, $\xi$ is a real $M \times 1$ vector, and $\xi^t \mathrm{g}$ denotes a scalar product.

In the case of a random process $f(\mathrm{r})$, each sample function corresponds to a vector $\mathrm{f}$ in an infinite-dimensional Hilbert space, so the frequency vector $\xi$ in (8.26) must be replaced by an infinite-dimensional vector $\mathrm{s}$ in the same Hilbert space as $\mathrm{f}$. That means that $\mathrm{s}$ describes a function $s(\mathrm{r})$, so the characteristic function becomes a characteristic *functional* $\Psi_{\mathrm{f}}\{s(\mathrm{r})\}$ or $\Psi_{\mathrm{f}}(\mathrm{s})$ for short. It is defined by

$$\Psi_{\mathrm{f}}(\mathrm{s}) = \langle \exp[-2\pi i (\mathrm{s}, \mathrm{f})] \rangle , \qquad (8.94)$$

where $(\mathrm{s}, \mathrm{f})$ is the usual $\mathbb{L}_2$ scalar product.

Equation numbers refer to Barrett and Myers, Foundations of Image Science

# Random processes for which the characteristic functional is known analytically

- Lumpy backgrounds (often used to simulate tissue inhomogeneity in medical imaging)
- Clustered lumpy backgrounds (good model for mammograms)
- Any Poisson random process (stationary or not)
- Any Gaussian random process (complex or real, arbitrary covariance)
- Fully developed speckle
- Certain texture models related to non-Gaussian speckle
- Image-plane irradiance in adaptive optics
- Any of the above after detection by a discrete detector array with Poisson on Gaussian noise

# Uses of characteristic functionals

- If the characteristic functional for the object random process is known, we can derive:
  - Any desired marginal PDF on the object
  - Any moment or covariance function of the object
  - The characteristic function for the image data through any linear -- and some nonlinear -- imaging systems
  - The object and noise terms in the data covariance matrix needed for Hotelling and Wiener observers
  - Ideal observer performance

Details in Barrett and Myers