SAG19: contract metrics for internal coronagraphs

D. Mawet, 11 June 2016

Literature on signal detection and speckle statistics

- Perrin, M. D., Sivaramakrishnan, A., Makidon, R. B., Oppenheimer, B. R., & Graham, J. R. 2003, ApJ, 596, 702
- Fitzgerald, M. P., & Graham, J. R. 2006, ApJ, 637, 541
- **Kasdin, N. J., & Braems, I. 2006, ApJ, 646, 1260**
- Soummer, R., Ferrari, A., Aime, C., & Jolissaint, L. 2007, ApJ, 669, 642
- Hinkley, S., Oppenheimer, B. R., Soummer, R., Sivaramakrishnan, A., Roberts, Jr., L. C., Kuhn, J., Makidon, R. B., Perrin, M. D., Lloyd, J. P., Kratter, K., & Brenner, D. 2007, ApJ, 654, 633
- Marois, C., Lafreniere, D., Macintosh, B., & Doyon, R. 2008, ApJ, 673, 647
- Mawet, D., Milli, J., Wahhaj, Z., et al. 2014, ApJ, 792, 97
- and many more

Contrast Definitions for Direct Imaging of Extrasolar Planets

John Krist (Jet Propulsion Laboratory, California Inst. of Technology)
26 May 2016

Background

There is the unambiguous, no-qualifier-prefix required definition of contrast: the ratio of the planet intensity to the star intensity when both objects are observed in exactly the same manner. When contrast is used to describe coronagraphic performance, however, contrast can be misused or misunderstood. Background contrast and detection limit contrast are two different things, and they are related to, but not equivalent to, intensity.

The field of interest (e.g., a coronagraphic dark hole) as measured with the detector will be filled with diffracted light from the star (mostly reduced, though not completely, by the starlight suppression system), scattered light speckles and stray light from optical defects (perhaps reduced by wavefront control), and extended astrophysical sources (e.g., zodi or exozodi). Even with a perfect detector and post-processing capabilities, the shot noise from these sets a limit on planet detectability. We assume here that the scattered & stray light and the astrophysical backgrounds are generally (but not completely) independent of the suppression technique, so we ignore them since we are concerned about comparative performances.

Community outreach: Exoplanet Imaging Facebook group (677 members)

+ Email contacts from

R. Belikov (NASA AMES),

S. Bryson (NASA AMES),

J. Mazoyer/L. Pueyo (STScI)

Thayne Currie, John Debes and 15 others

Dimitri Mawet Ok, I think I need to be more provocative, and troll this thread to initiate a constructive discussion. Our current contrast and SNR definitions are wrong and meaningless. Implications are. Still too many false positives (especially confusion with exte... See More

Like · Reply · 8 · June 6 at 8:40am

Michael Meyer Can we agree to separate the discussion between false alarm probabilities (which we ought to be able to agree on to interpret null results and vital for surveys) versus quoting measured parameters for detections which are fundamental to characterization (and in my opinion trickier)?

Like · Reply · 1 · June 6 at 2:04pm

Dimitri Mawet First input! Yes, thank you!

Like · Reply · June 6 at 2:40pm

Brendan Bowler What constitutes a detection? My impression is that most groups apply a range of conservative and aggressive reductions to each dataset, then search for candidates visually or (ideally) with point source finding algorithms. As far as contrast curves,... See More

Like · Reply · 3 · June 6 at 2:53pm

Angelle Tanner Im adding myself since I'm interested

Like · Reply · June 6 at 2:56pm

Matthew Kenworthy How do you build up your noise distribution at small angles, when it appears that your distribution changes with different observing conditions? Can this be a mappable/determinable function, or are you restricted to using the data from that particular observation?

Like · Reply · June 6 at 2:57pm

Katie Morzinski Could you also email Forum HRA? Not everyone is on FB.

Like · Reply · June 6 at 5:38pm

Julien Girard Shall noise calculation boxes (size) and small angles penalties be determined in function of the data resolution elements (Airy for space and super xAO but FWHM for partial AO correction on the ground)? Should there be a common "rule" for that?

Like · Reply · June 7 at 6:41am

Jason Wang Building off what Brendan Bowler said, to determine detection limits, you also have to determine what is your false alarm probability (i.e. how many speckles do you want to follow up) to determine how conservative/aggressive your detection threshold is. You can go deeper, but also have to accept more false positives.

Like · Reply · 1 · June 7 at 11:53am

Andy Skemer Deep observations spanning several days will have to contend with orbital motion. "De-orbiting", leads to more false-alarms.
http://arxiv.org/abs/1304.5853

[1304.5853] Direct Imaging in the Habitable Zone and the Problem of Orbital Motion

ARXIV.ORG

Like · Reply · June 7 at 1:19pm

Angelle Tanner Is there a desire for some type of small study where people evaluate the same data set(s) using their chosen techniques and we compare results?

Like · Reply · June 7 at 2:33pm

Dimitri Mawet Angelle Tanner, I think this is a great idea. We could agree on a famous public data set (beta Pic, HR8799, etc), and ask volunteers to run their preferred pipelines, quote astrometric and photometric values for the planet(s), uncertainties and provide contrast maps or curves. The comparison of results would be very informative.

Like · Reply · 1 · June 7 at 2:50pm

Julien Girard Yes, this is the principle of the beauty contest at the SPIE where the long baseline (optical) interferometry community is invited to reconstruct images out of data sets (simulated, modified VLTI or CHARA datasets). It would be time to organize somethi... See More

Like · Reply · June 7 at 3:10pm

> **Angelle Tanner** Can we not call it a beauty contest?
>
> Like · Reply · 2 · June 7 at 3:12pm
>
> **Julien Girard** We can call it blind planets or beauty contest. We can have a contest to find a name 😜
>
> Like · Reply · June 7 at 4:24pm
>
> Write a reply...

John Debes I would recommend both a space based and ground-based dataset. The Fomalhaut datasets are well suited from space.

Like · Reply · June 7 at 3:57pm

Christian Marois Using famous data sets (HR8799) would be good, but I think adding a bunch of simulated planets of known properties in "what ever" data sets would be more interesting IMO (no one knows the real positions/spectra of the HR8799 planets, so how are we goin... See More

Like · Reply · June 7 at 6:34pm

Angelle Tanner I assume we are interested in both the ability to detect a plane (or to say a planet is "detected")t as well as a common method for estimating detection limits when there is no planet. I've seen people threat these two issues separately.

Like · Reply · June 7 at 8:57pm

↳ Angelle Tanner replied · 2 Replies

Motivations for SAG19

- There is confusion in the contrast metrics terminology and definitions
- Discrepancies between what the community (95% ground-based) uses and what coronagraph designers/builders take as a metric
- There is no gold standard in the community either
- Every group, individual has their own way of measuring “contrast”, and perhaps worse, assess error bars (astrometry, photometry)
- Leads to erroneous/ambiguous results when compiling and comparing published data (occurrence rates, orbital analysis, etc.)

The need for a standardization of high contrast imaging metrics across the board

- We need to go back to the basics of signal detection theory, and take example at the medical imaging world where definitions and rigorous methodologies are a matter of “life and death”

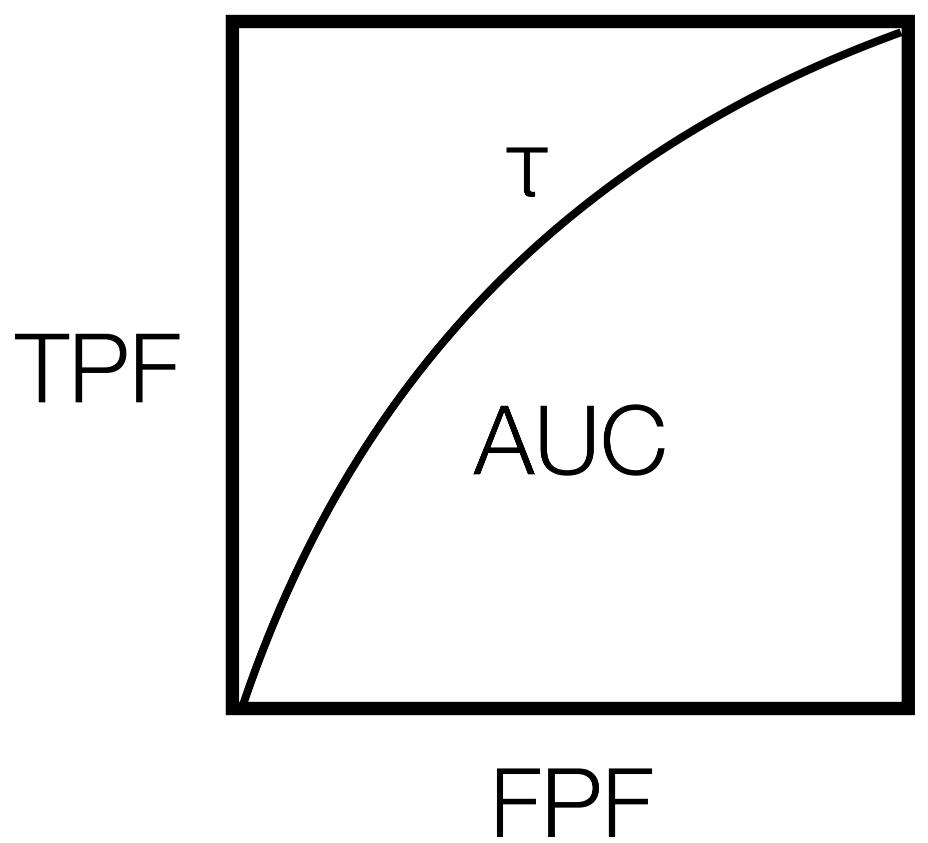
- The Receiver Operating Characteristic (ROC) curve, plotting the true positive fraction as a function false negative fraction, as well as the Area Under the Curve (AUC) is the gold standard

- ROC/AUC formalisms automatically captures false alarm probabilities and completeness

Signal Detection Theory: receiver operating characteristics

	H_1: signal present	H_0: signal absent
Detection	True Positive	False Positive ***type I error***
Null result	False Negative ***type II error***	True Negative
	TPF = sensitivity = power = TP/(TP+FN)	**FPF = 1-specificity = 1-CL = FP/(FP+TN)**

ROC curve



$$TPF = \frac{TP}{TP+FN} = \int_{\tau}^{+\infty} pr(x|H_1)dx$$

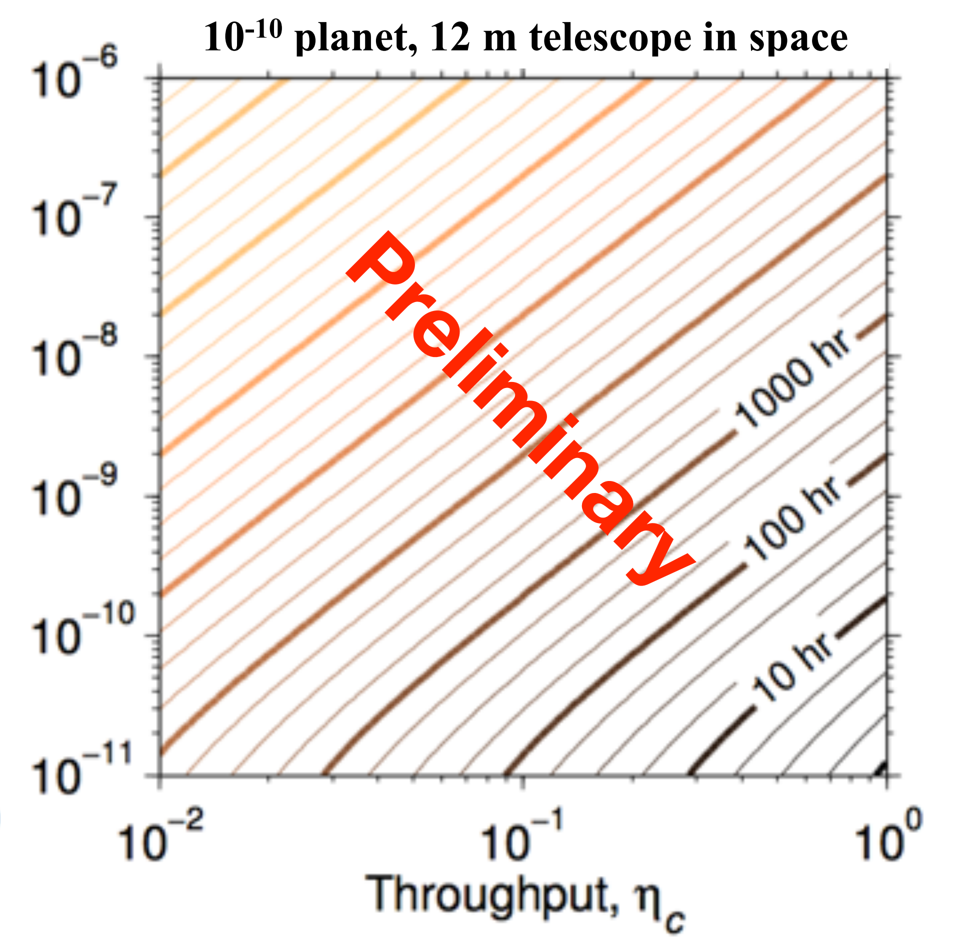
$$FPF = \frac{FP}{TN+FP} = \int_{\tau}^{+\infty} pr(x|H_0)dx$$

Terminology

- **Contrast**: ratio of physical quantities, planet / starlight brightness
- **Signal**: mean intensity of planet signal integrated in defined aperture (~FWHM) at the location of the planet, accounts for instrument losses (encircled energy ~ Strehl ratio)
- **Noise**:
  - Noise sources at the pixel locations of the planet (photon noise from planet and residual starlight, background, zodi, readout noise, dark current, …)
  - Speckle noise:
    - Underlying the planet signal => affects characterization
    - Adjacent to the planet => affects detection
      - Adjacent:
        - spatially and temporally (equivalent only if noise is ergodic)
        - annulus or arc at planet radius
        - box centered at the location of the planet

Fundamental SNR equation and implications

- SNR = η P / √(S/c + D)
  - P=planet light
  - η=throughput within aperture
  - c=starlight suppression factor at planet location
  - S=stellar noise residual w/o coronagraph
  - D=detector, background, zodi noise
- SNR # η . √c
- ***Gaining a factor 10 in contrast is useless if throughput is reduced by a factor 3***



Ruane et al. 2016, in preparation

SNR alone has no direct statistical meaning

- SNR of 5 does not automatically imply false positive fraction of 3e-7
- Strongly depends on underlying noise statistics
- Speckle noise is NOT Gaussian
- Adjustments can be extreme, and have huge consequences on performance metrics

Knowing your underlying speckle noise distribution

Modified Rician:

$$p_{MR}(I, I_c, I_s) = \frac{1}{I_s}\ \exp\left(-\frac{I+I_c}{I_s}\right)\ I_o\left(\frac{2\sqrt{II_c}}{I_s}\right)$$

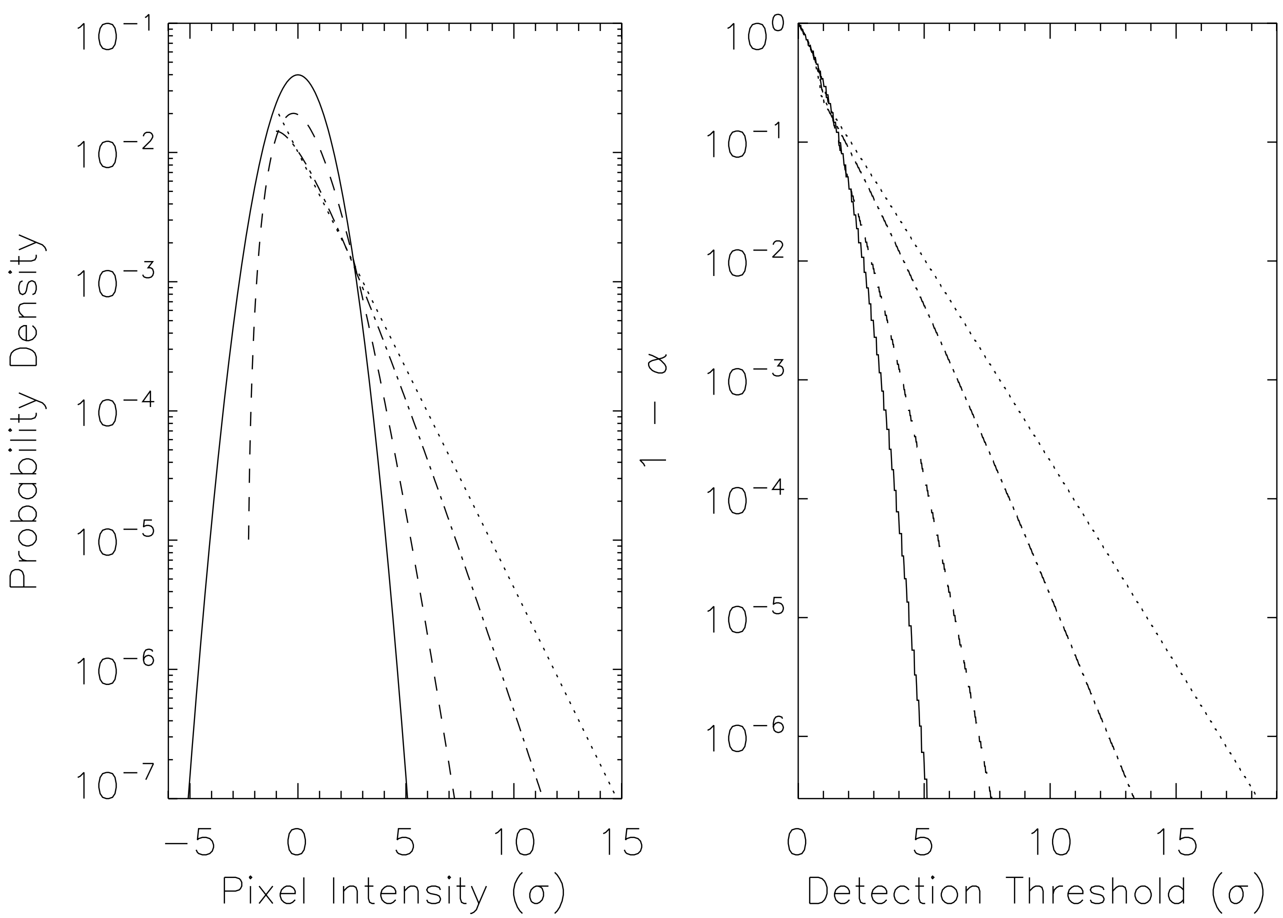


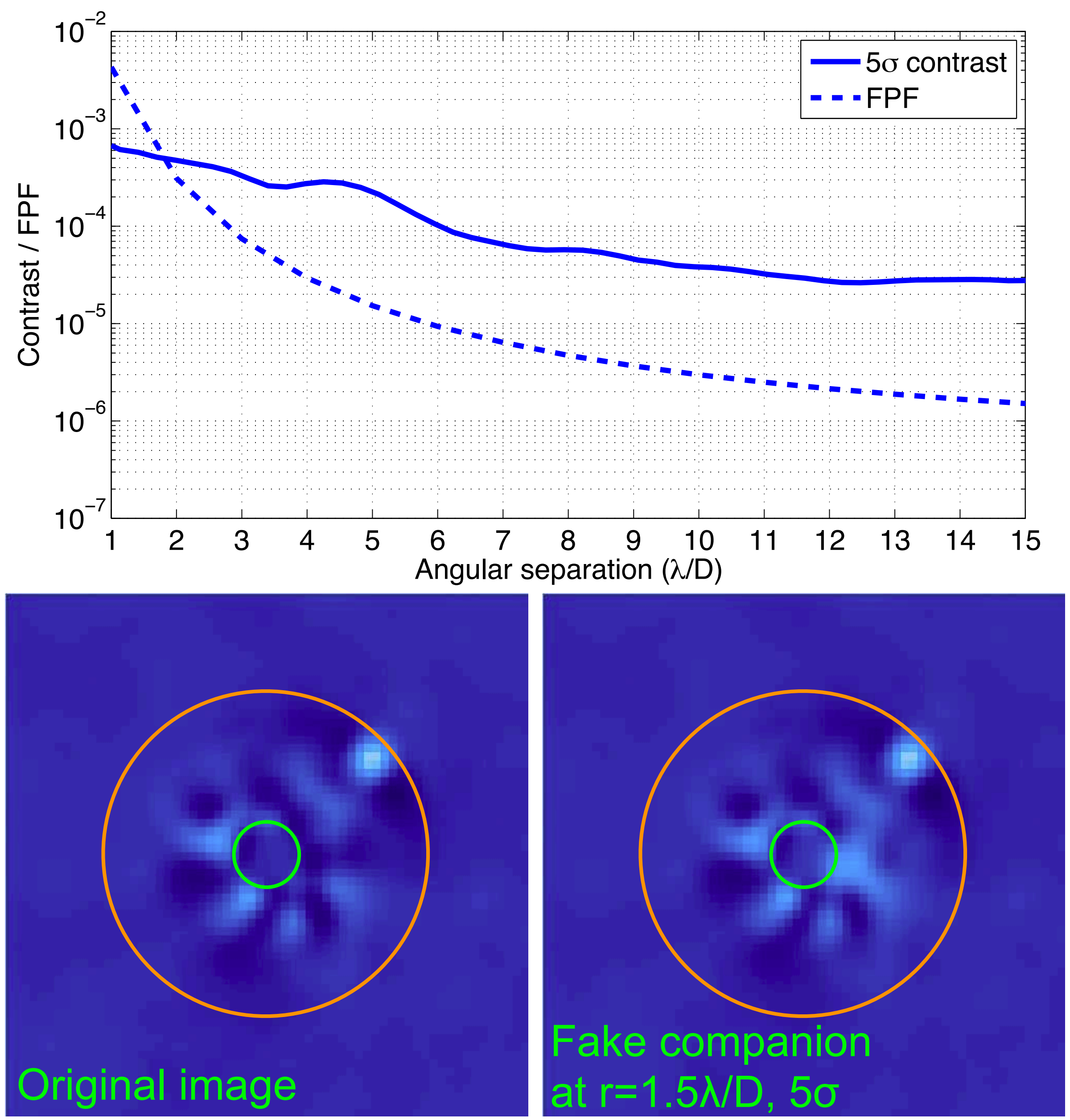
TABLE 1

CL Extrapolation Accuracy for Simulated Statistical Distributions

Statistics	No. Resolution Elements	Expected d (σ)	$\langle d \rangle$ (σ)	Standard Deviation (d) (σ)
Gaussian............	10^3	5.0	5.4	1.1
	10^4		5.33	0.40
	10^5		5.06	0.29
	10^6		5.06	0.11
MR10	10^3	7.7	9.3	1.9
	10^4		8.20	0.95
	10^5		8.02	0.59
	10^6		8.02	0.66
MR1	10^3	13.5	14.9	2.5
	10^4		13.95	0.98
	10^5		13.44	0.50
	10^6		13.48	0.37
MR01	10^3	18.2	18.7	2.6
	10^4		17.2	1.4
	10^5		17.13	0.40
	10^6		17.18	0.41

Marois et al. 2008

What happens at small IWA?

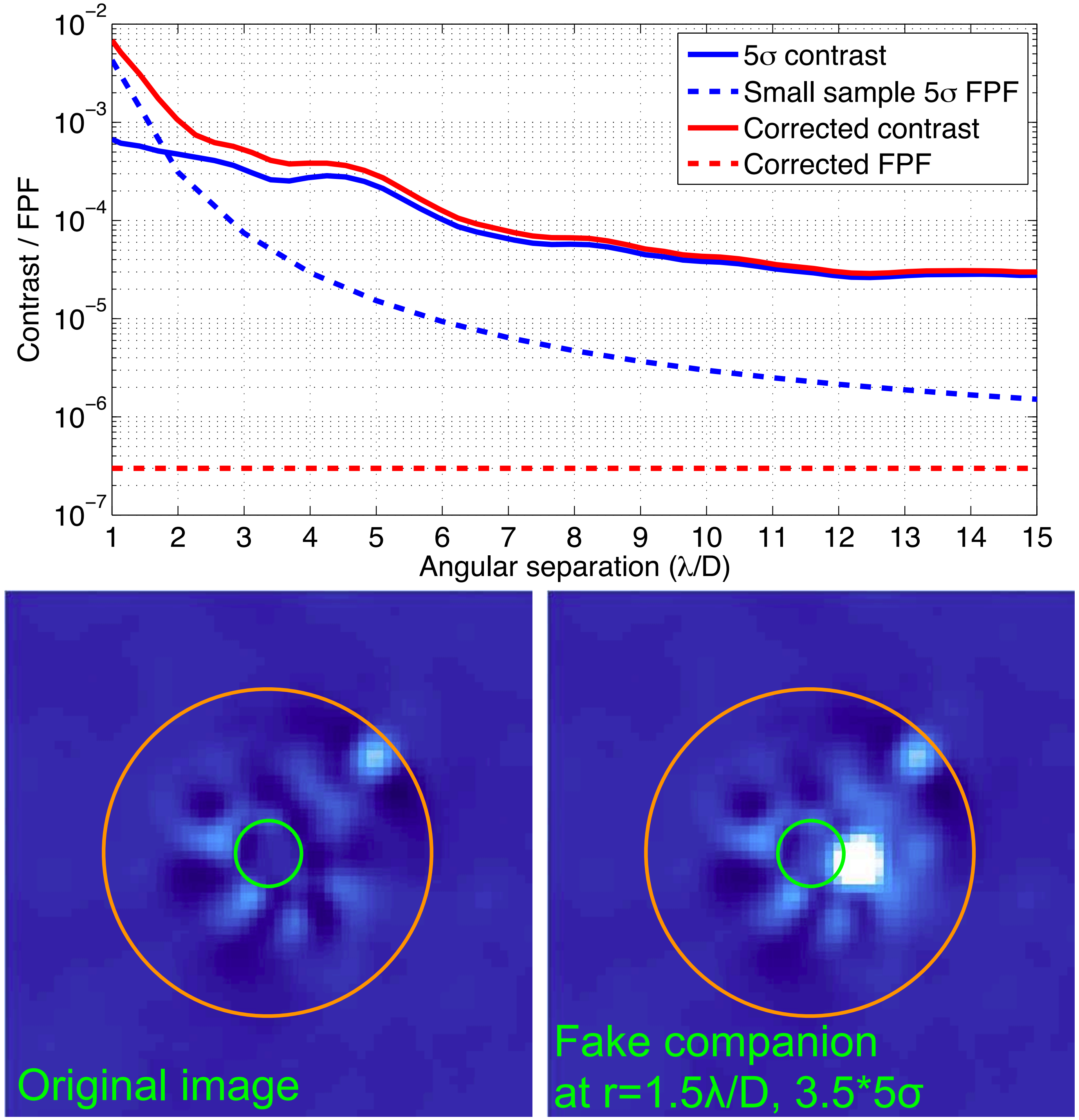


σ, here actually is the empirical standard deviation

The true STD of the underlying distribution is unknown!

Mawet et al. 2014

After small sample statistics Student t correction

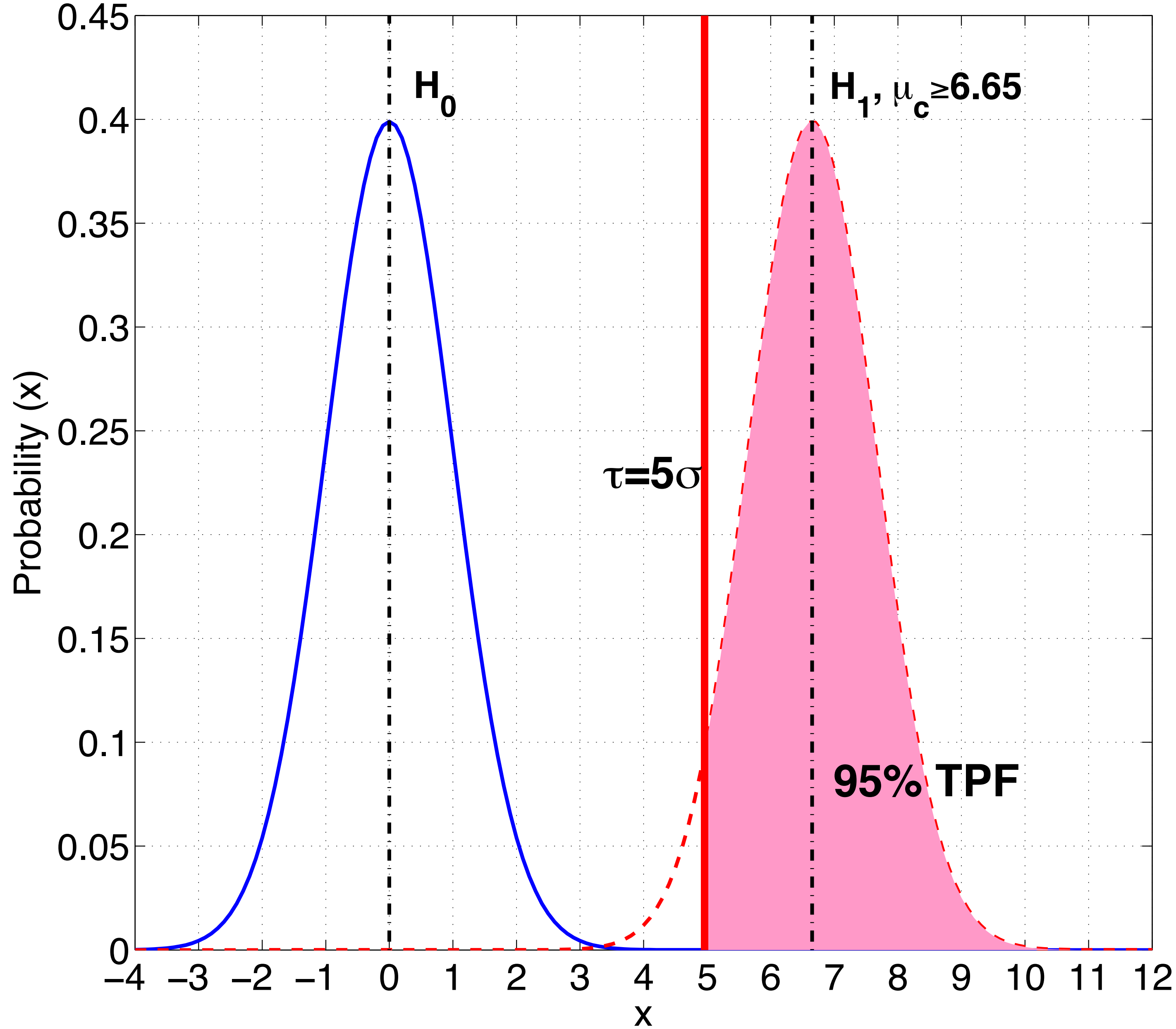


Mawet et al. 2014

Choice of threshold for detection is arbitrary and not properly motivated

- What do we want?
    - Minimize FPF? maximize TPF?
    - Both => maximize AUC (e.g. Hotelling observer)
- False positives: does 3e-7 FPF make sense? How many false alarm are we ready to tolerate?
- True positive fraction (recovery rate, or completeness): e.g., at $\tau = 5$, what is the TPF of $P = \tau\,\sigma = 5\,\sigma$? Assuming Gaussian noise (which is wrong).

Ideal Gaussian case



$$FPF = \int_{\tau}^{+\infty} pr(x|H_0)dx = \int_{\tau}^{+\infty} \mathcal{N}(\mu, \sigma^2)dx$$

where

$$\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\int_{\mu_c - \tau}^{+\infty} pr(x|H_1, \mu_c)dx = TPF$$

Null results and detection limits

- Coronagraph and WFC impact planet throughput (SNR # η . √c)
- Post-processing too!!! And in major ways (sometimes >>10x)
- Requires injecting fake planets in raw data
  (standard procedure in Kepler, see Petigura et al. 2013)
- Adopted by the ground-based high contrast imaging community for at least half a decade
  - Scanning location required to sample recovery rate as a function of location in the image
  - Injected fake planets need to account for encircled energy variations as a function of location

Detection: confidence level, characterization, astrometric and photometric error bars

- Detection above threshold? What is the likelihood of false positive? The need for follow-up?
- Characterization using **forward modeling**:
  - Fake companion (gold standard), and MCMC exploration over x,y,f => error bars
  - Perturbation based technique. Proper to PCA post-processing. Similar results as the fake companion technique but faster. See Pueyo 2016.

Coronagraph design metrics

- **Starlight suppression (mean intensity)**:
  - Integrated in the pupil
  - Integrated in the focal plane from r_{min} to r_{max}
  - At a specific location in the image plane (box of arbitrary shape)
  - fraught with potential interpretation errors, and cannot be used across different instrument
- **Inner working angles (IWA):**
  - 50% off-axis throughput (which throughput?)
- **Outer working angle (OWA):**
  - Field of view? Edge of the control area?
- **Throughput**:
  - Telescope + instrument (optics)
  - Coronagraph throughput: integrated over aperture
  - End-to-end integrated over aperture is only what matters
- **Bandwidth and spectral resolution**
- So far not concerned at all by speckle statistics
  - speckle control affects speckle statistics, so does post-processing

Plan for SAG19

- Seek community volunteers and co-chair: on-going
  - setup website, sharing tools
- Draft detection metrics charter with subset of volunteers
- Broadcast it widely and seek feedback from community
- Initiate data challenge based on real data from ground-based telescopes, and space-based one (HST + WFIRST simulations/lab data?, interface to WFIRST SITs)
- Write final report and recommendations to, in particular, TDEM

2016 Sagan Exoplanet Summer Workshop

Is There a Planet in My Data? Statistical Approaches to Finding and Characterizing Planets in Astronomical Data

JULY 18-22, 2016

HOSTED BY THE NASA EXOPLANET SCIENCE INSTITUTE

AT THE BECKMAN INSTITUTE AUDITORIUM, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA

IMPORTANT NOTE: DUE TO SPACE LIMITATIONS, THE NUMBER OF ATTENDEES AT THE WORKSHOP WILL BE CAPPED AT 175 AND THE NUMBER OF PARTICIPANTS IN THE HANDS-ON SESSIONS WILL BE CAPPED AT 125.

MAY 16 UPDATE: THE HANDS-ON SESSIONS ARE NOW FULL. IF YOU REGISTER AFTER TODAY AND WISH TO PARTICIPATE IN THE HANDS-ON SESSIONS, YOUR NAME WILL BE PLACED ON A WAITING LIST.

BOX LUNCHES AND TICKETS FOR THE WORKSHOP DINNER MAY NOW BE PRE-PURCHASED WITH A DEADLINE OF JULY 12. CLICK HERE FOR MORE INFORMATION AND THE PURCHASE INFORMATION.

The 2016 Sagan Summer Workshop will focus on data analysis techniques used to find planets in various types of data. In particular, leaders in the field will discuss Monte Carlo Markov Chain (MCMC) and Bayesian inference relevant to transit analysis and spectral retrieval as well as RV analysis. Image processing techniques such as Principal Component Analysis (PCA), LOCI, and KLIP methods will also be discussed. In addition, for each of these areas, noise sources and mitigation strategies will be highlighted. Attendees will participate in hands-on group projects and will have the opportunity to present their own work through short presentations (research POPs) and posters.

There is no registration fee for the workshop. If you have registered and will no longer be able to attend, please let us know so that your spot may be given to someone else.

KECK INSTITUTE FOR SPACE STUDIES

Earth Planetary Keck Astrophysics Engineering

Generously supported by the W. M. Keck Foundation

- Home
- Visit
- Programs
- Events
  - Workshops
  - Lectures
  - Short Courses
  - Symposia
  - Special Events
- People
- Archive
- About
- Follow Us
- Calendar

Exoplanet Imaging and Characterization: Coherent Differential Imaging and Signal Detection Statistics

August 22 - 26, 2016
California Institute of Technology - Pasadena, CA 91125

Team Leads

Michael Fitzgerald
UCLA

Dimitri Mawet
Caltech

Gautam Vasisht
JPL

Description

Workshop Overview:

Direct imaging of a planet around another star is exceedingly challenging. For even the closest stars observed with the largest ground-based telescopes, the angular separation between star and planet will be near the classical diffraction limit. Moreover, a typical star will be about a billion times brighter than the planet to be imaged, a challenge even for the most stable telescope in space. The planetary image is also buried in "speckle noise," which is the result of uncorrected wavefront errors that propagate through the atmosphere and even the most accurately polished optical system. This speckle noise has complex properties which are different from planetary signals. While algorithms now exist that exploit some differences between the signal and noise characteristics, there has been little effort to address the full problem in a rigorous and comprehensive way. Our focus on speckle discrimination and control is motivated by key scientific measurements of exoplanetary systems:

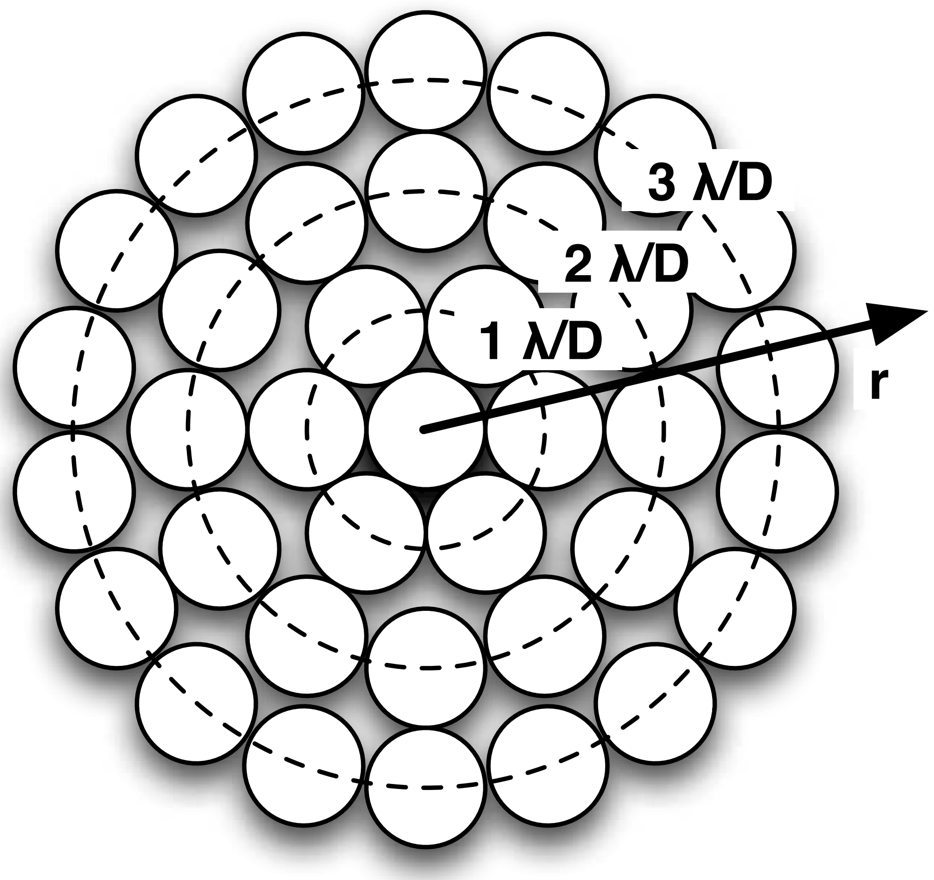
- Pure detection: Is there a planet present in the image(s)?
- Astrometry: Where precisely is the planet located?
- Photometry: How bright is the planet, and does its brightness vary with time?
- Estimation of orbital parameters: How does it move relative to other bodies in the system?
- Spectrometry: What are the spectral characteristics of the light from the planet?
- Detection of life: Does the spectrum contain components consistent with living organisms?

Given the advent of high-precision focal plane wavefront control and low-noise fast-frame-rate detectors as well as the ongoing development of new facilities for exoplanet study, our proposed KISS workshop seeks to address several questions related to the development of statistically grounded strategies for detecting faint signals in the presence of both coherent and incoherent backgrounds:

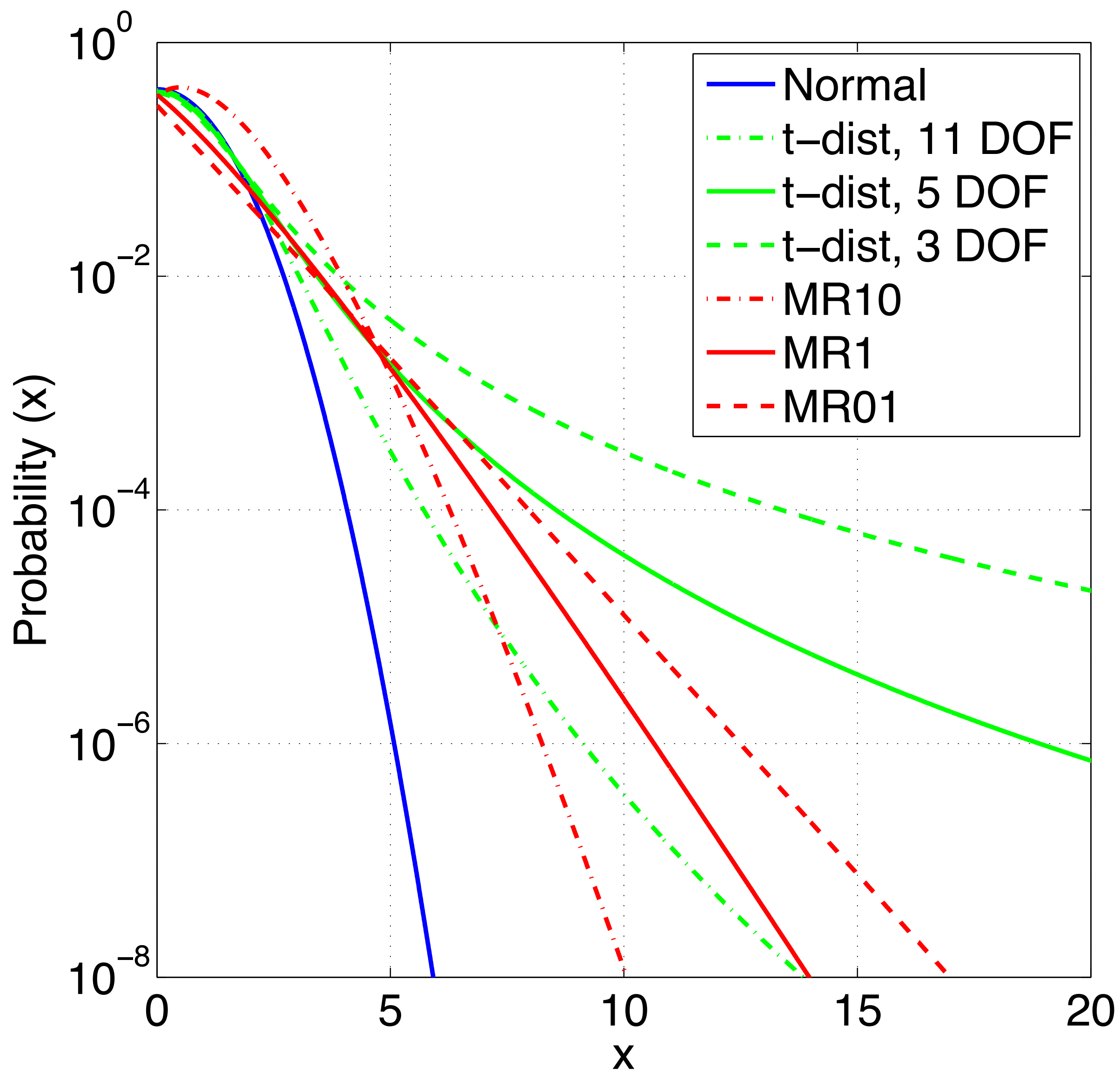
- What are the fundamental limits to focal-plane wavefront sensing and coherent differential imaging?
- How is modulation best used in source-speckle discrimination and in speckle control?
- How do ground-based focal-plane wavefront sensing and control differ from the space-based case?
- What is the potential impact of new post-processing techniques, and rigorous statistical analyses on the next-generation instruments for extremely large ground- and space-based telescopes?

Backup

Knowing your underlying speckle noise distribution



DOF getting smaller and smaller



$$p_t(x,\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

New definition of SNR and contrast

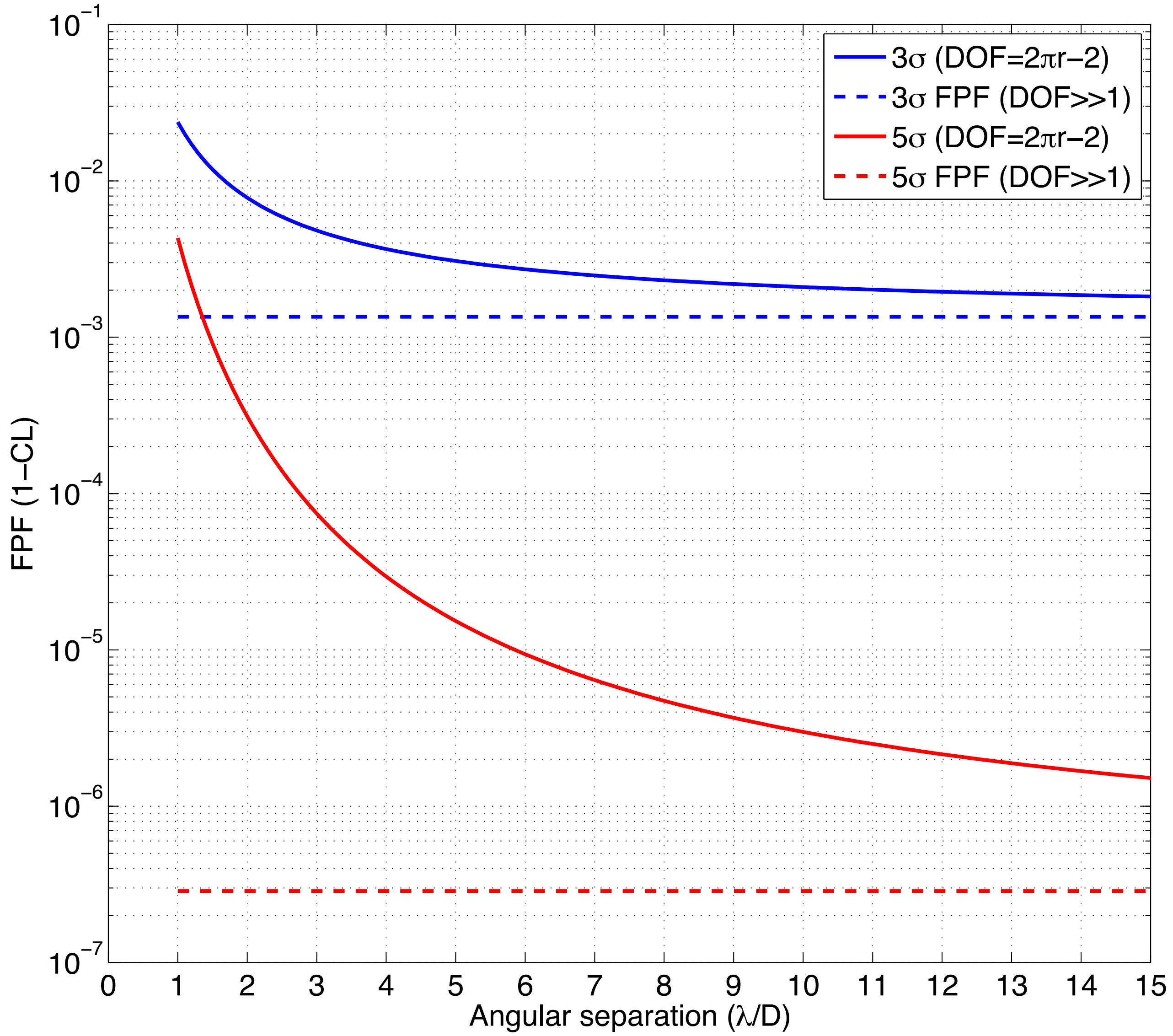
$$p_t(x, n_2 - 1) \equiv \frac{\bar{x}_1 - \bar{x}_2}{s_2\sqrt{1 + \frac{1}{n_2}}}$$

x_1 = test speckle intensity
x_2 = average of remaining n_2 speckle intensities at r
s2 = pooled standard deviation over n_2 remaining speckles at r

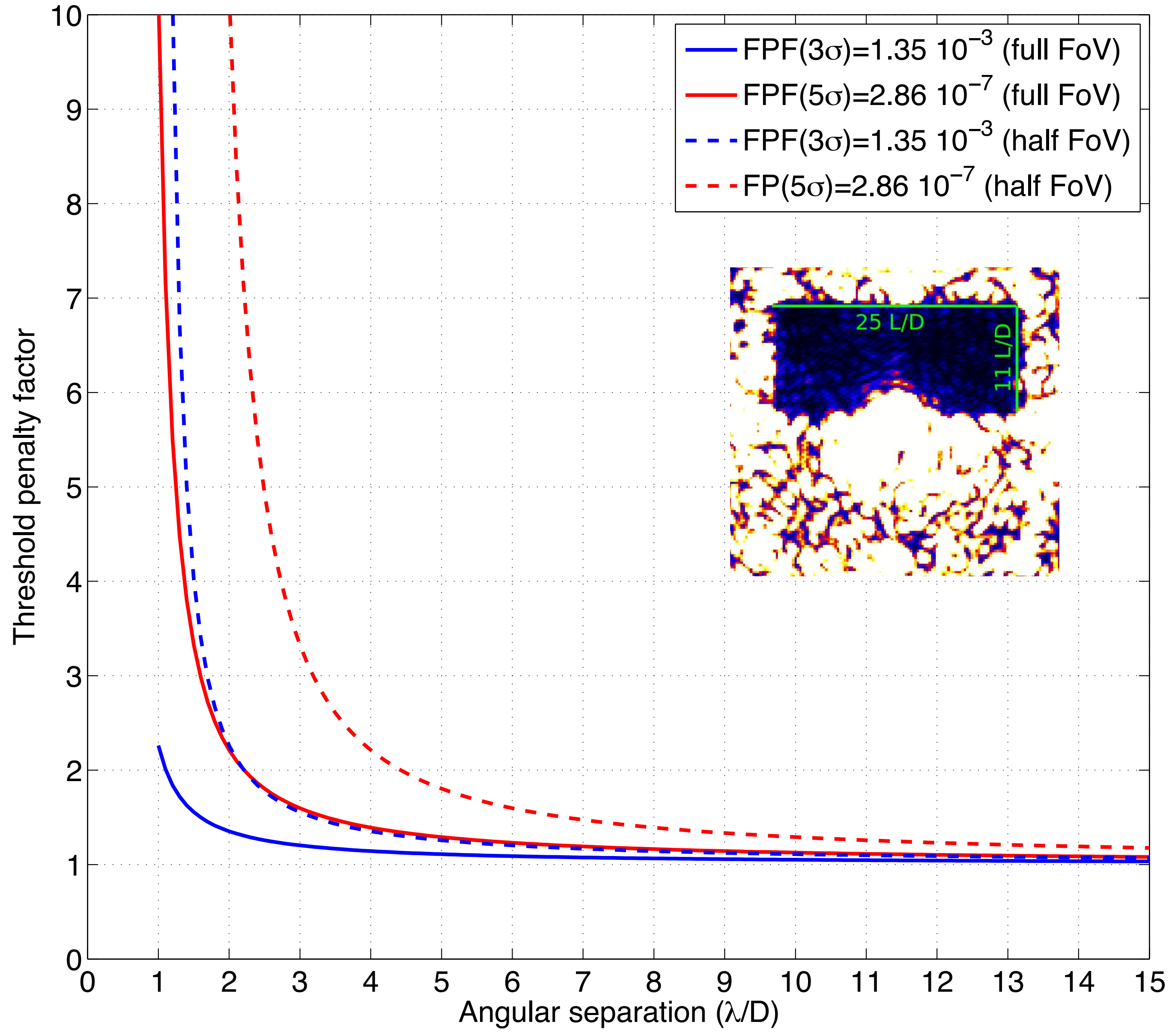
SNR is meaningless without corresponding FPF (1-CL)

$$FPF = \int_{\tau}^{+\infty} pr(x|H_0)dx = \int_{\tau}^{+\infty} p_t(x, n_2 - 1)dx$$

Consequences in terms of FPF



Consequences in terms of threshold and thus contrast



What if underlying noise is not Gaussian