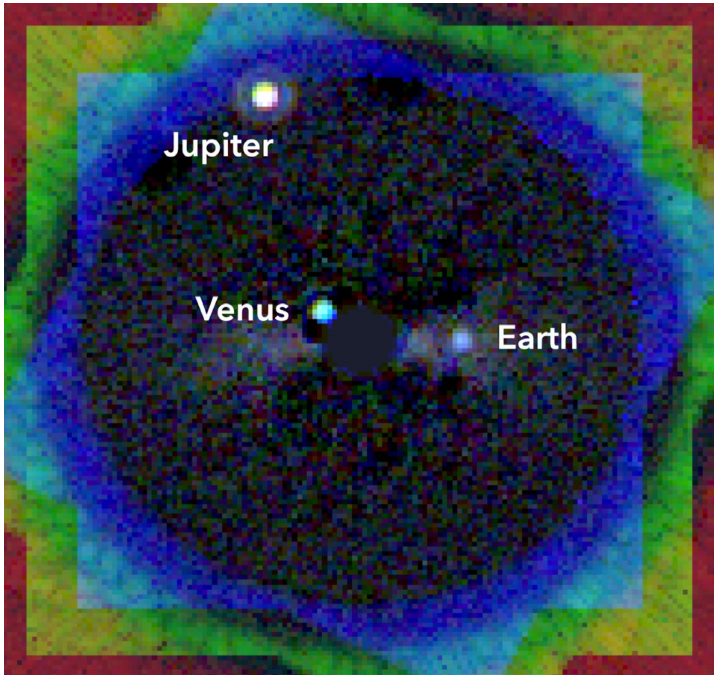
Coronagraphy 101

N. Jeremy Kasdin
Eugene Higgins Prof. of Mechanical & Aerospace Engineering, Emeritus
Princeton University

August 1, 2023



Acknowledgments

I am indebted to my direct imaging and Roman colleagues for their help and contributions to this talk:

Vanessa Bailey, Bruce Macintosh, Maggie Turnbull, Feng Zhao, Ilya Poberezhkiy, Rick Demers, Eric Cady, A J Riggs, Bijan Nemati, Bertrand Menneson, John Trauger, John Krist, Fang Shi, Pin Chen, and the entire CGI team.

Neil Zimmerman, Tyler Groff, Missy Vess, and the team at Goddard Space Flight Center.

Laurent Pueyo, Remi Soummer, Roser Juanola-Parramon, and Chris Stark at STScI.

Gary Ruane (Caltech), Dmitry Savransky (Cornell), Aki Roberge (GSFC), Leslie Rogers (Chicago), Courtney Dressing (Berkeley).

My students and postdocs: Christian Delcroix, Jessica Gersh-Range, Anthony Harness, Leonel Palacios Moreno, Mia Hu, Leonid Pogorelyuki, He Sun.

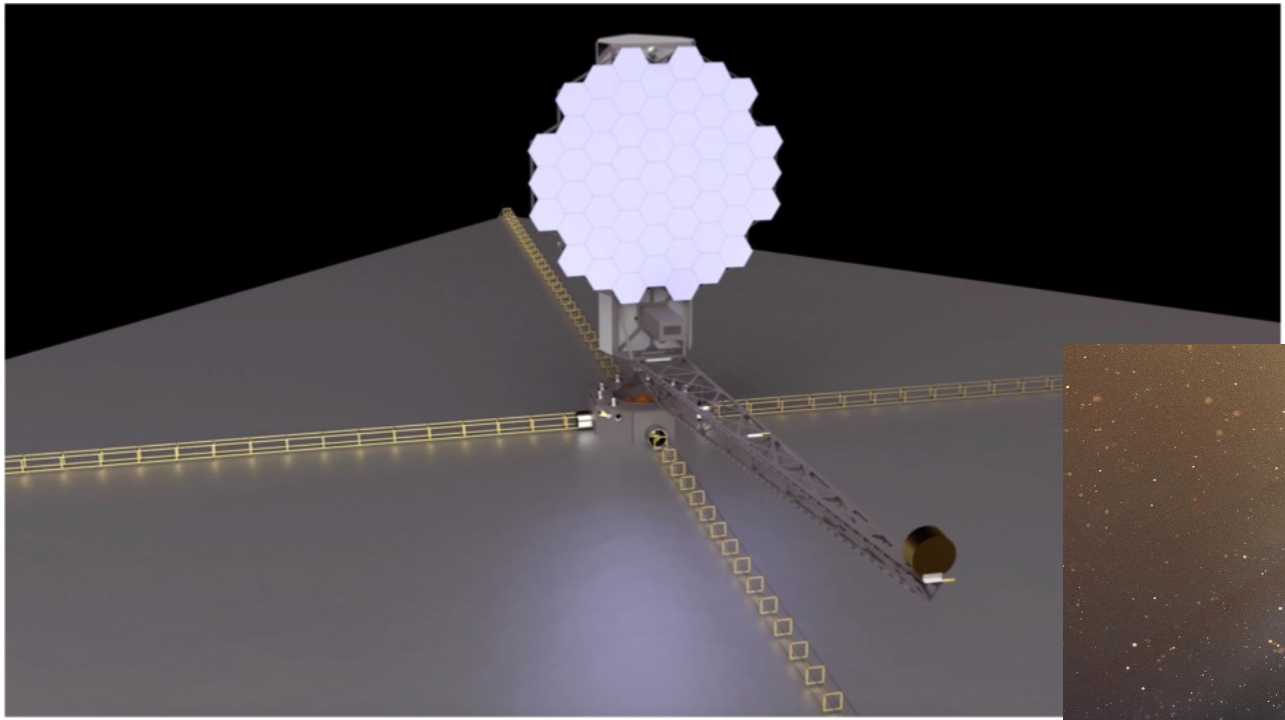
Astro2020

Pathways to Discovery in Astronomy and Astrophysics for the 2020s

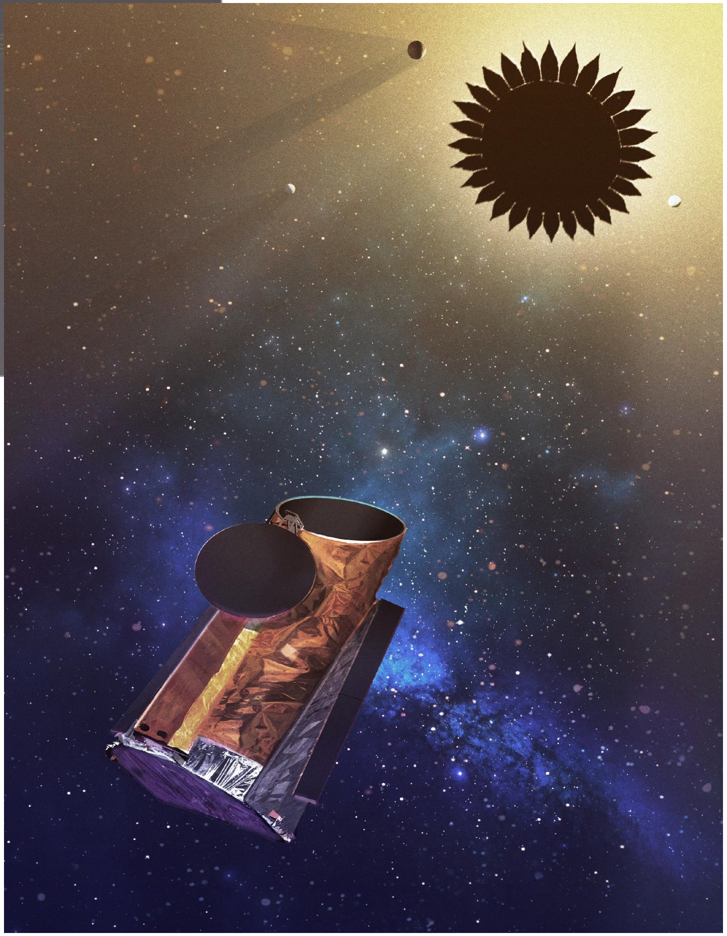
The decadal survey recommends a large (~6m diameter) Infrared/Optical/Ultraviolet space telescope with high-contrast imaging and spectroscopy as the first mission to enter the Great Observatories Mission and Technology Maturation Program. This is an ambitious mission with the goal of searching for biosignatures from habitable zone exoplanets and providing a powerful new facility for general astrophysics.

NASA has named this mission the Habitable Worlds Observatory (HWO)

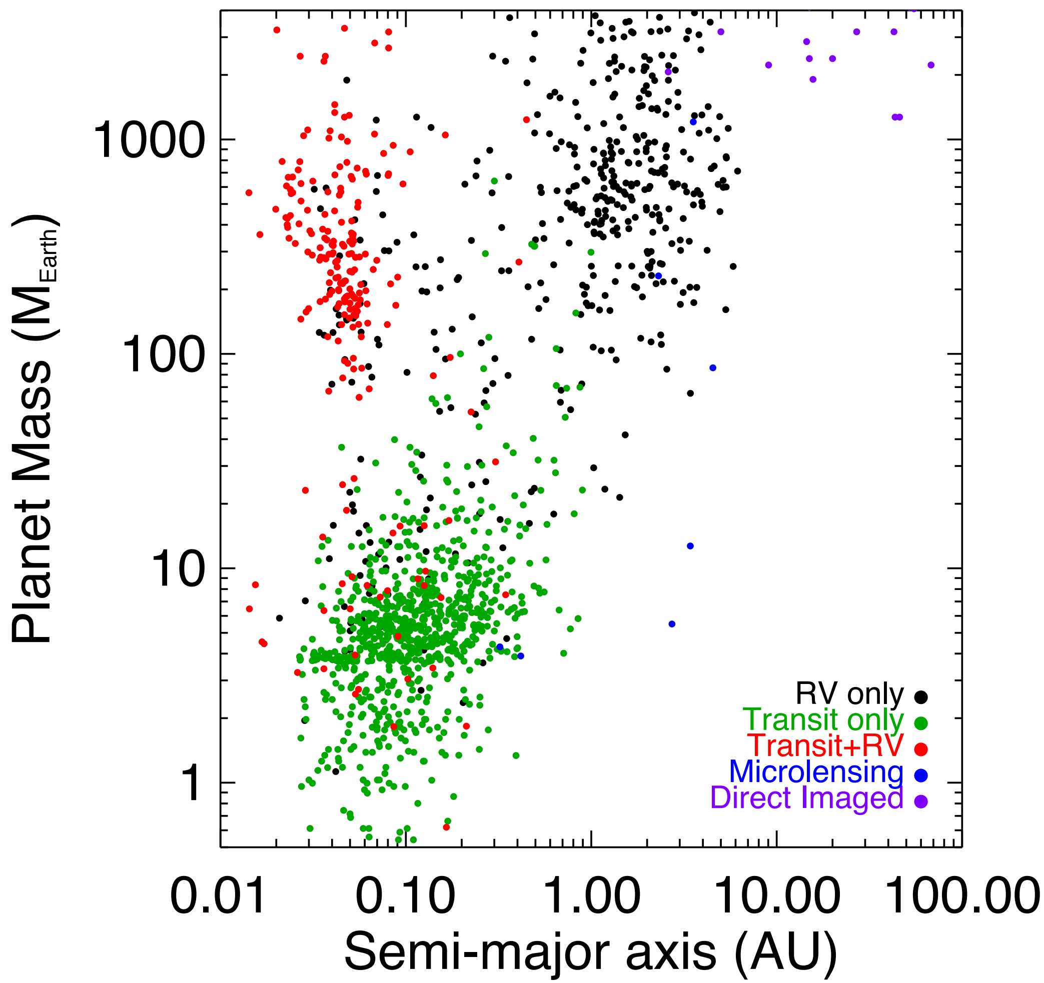
LUVOIR B



HabEx



Why Direct Imaging?



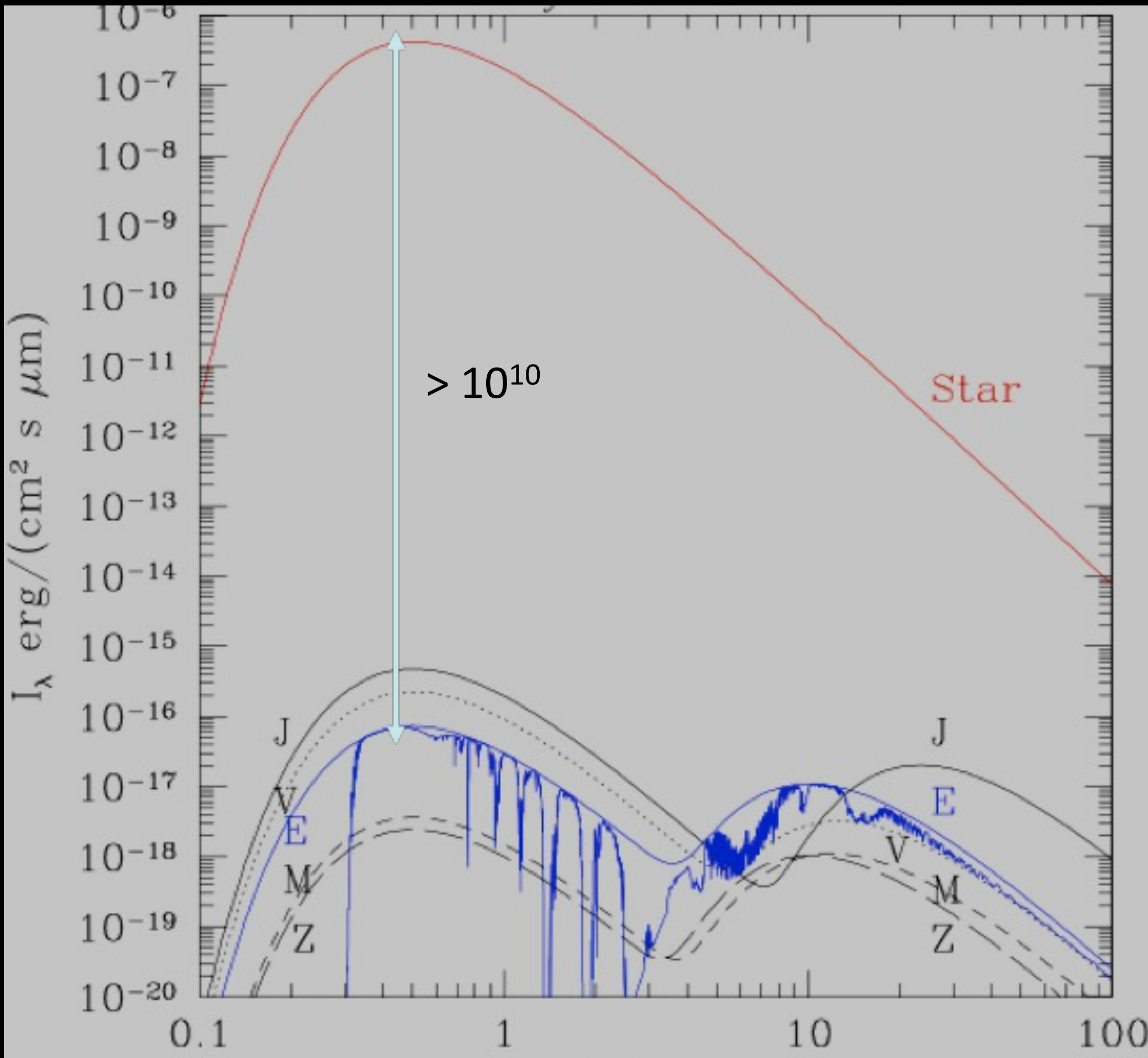
Why Direct Imaging?

- Statistical Properties – probing the outer parts of solar systems
- Detailed Characterization – determining the composition and detailed state of planetary atmospheres.
- Formation mechanisms – measuring parameters that constrain formation theories.
- Ultimately – determining whether life-bearing planets are common.
- Imaging is visually compelling – Great public interest

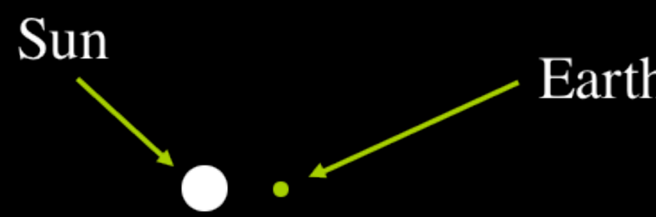
See colloquium by Giada Arney

Why is direct imaging a challenge?

Flux Ratio and Angle

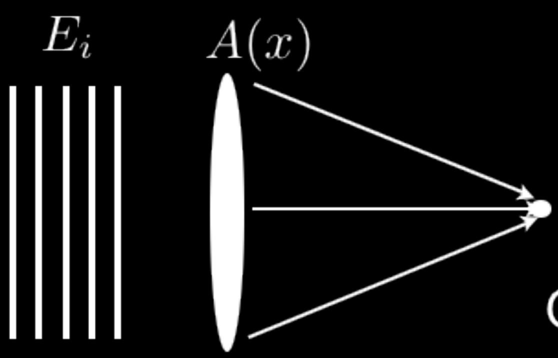


Traub & Jucks



At 10 pc distance, the angular separation between a planet in the habitable zone and its star is 100 marcsec

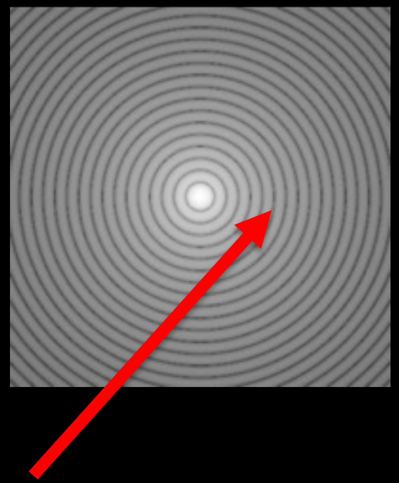
The Contrast Problem – Diffraction

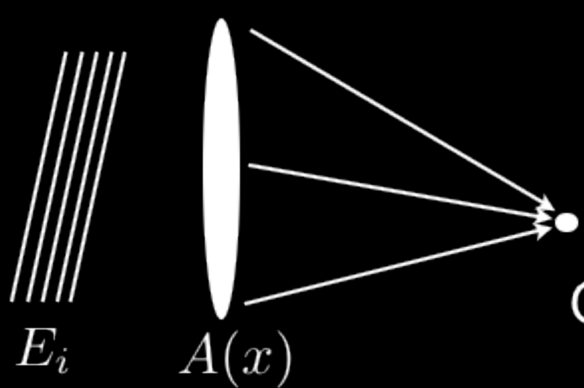


$$P_c(\omega) = |\mathcal{F}\{E_i A(x)\}|^2$$

On-Axis Point Spread Function

Circular Aperture

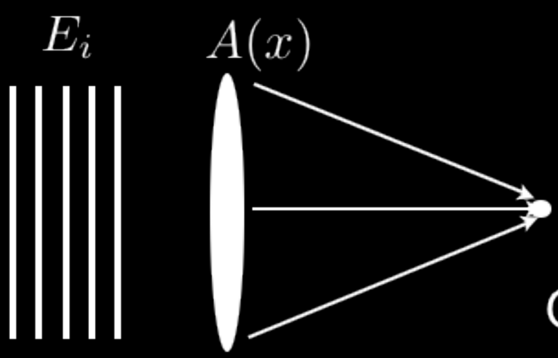




$$P_c(\omega) = |\mathcal{F}\{E_i e^{-x\theta/\lambda} A(x)\}|^2$$

Off-Axis Point Spread Function

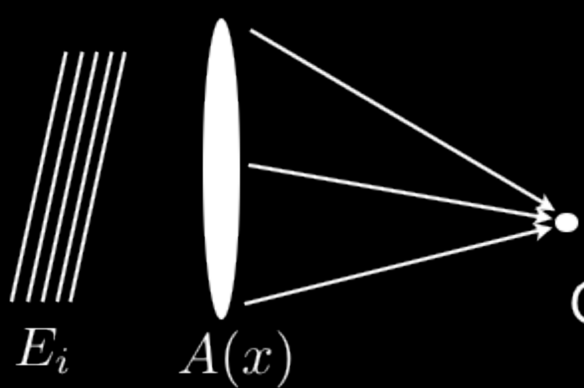
The Contrast Problem – Diffraction



$$P_c(\omega) = |\mathcal{F}\{E_i A(x)\}|^2$$

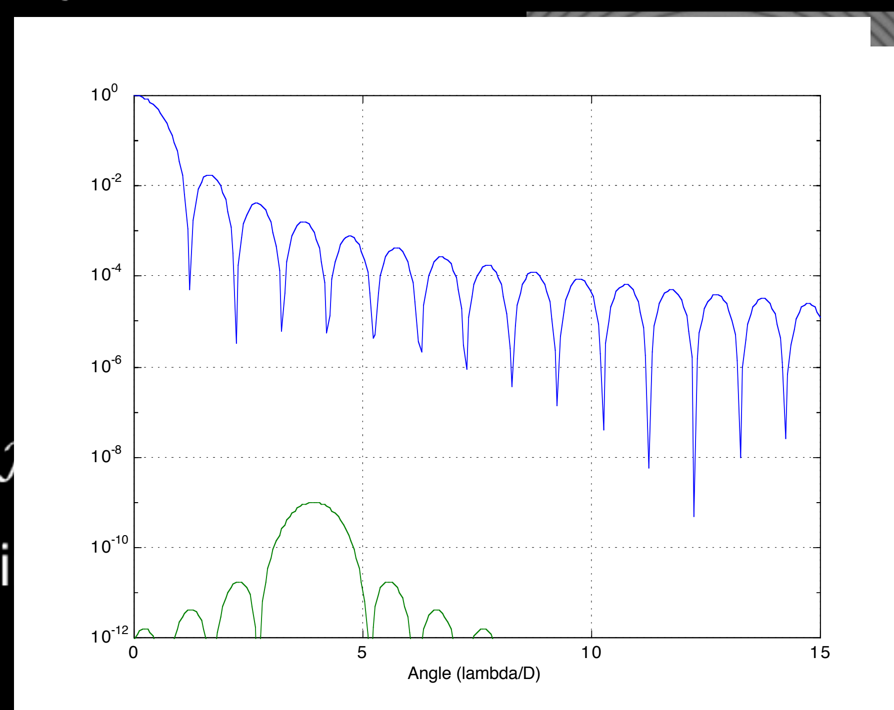
On-Axis Point Spread Function

Circular Aperture



$P_c(\omega) = |\mathcal{F}$

Off-Axis Poi

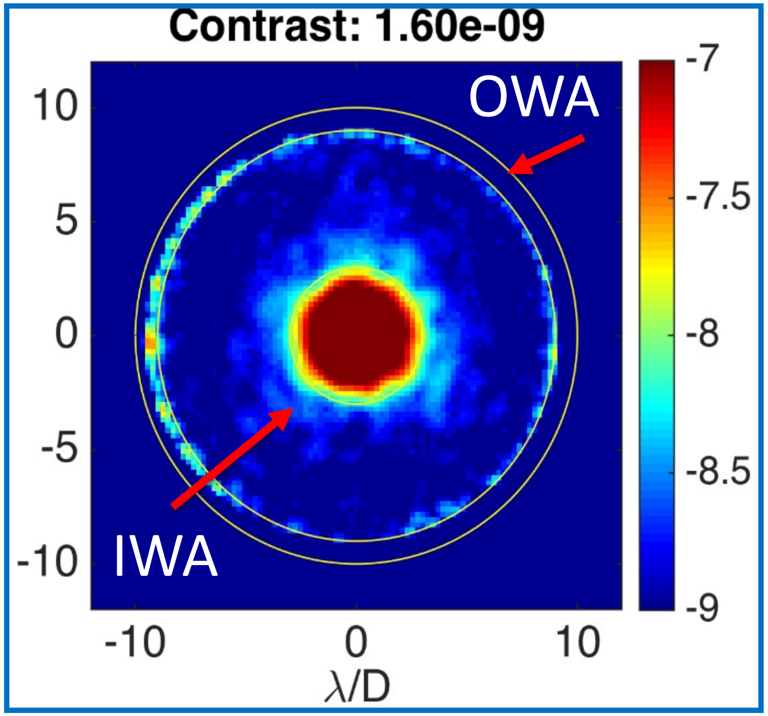


To image the planet, we must create *high contrast* in the final image plane, lowering the stellar halo to at or below the peak intensity of the planet.

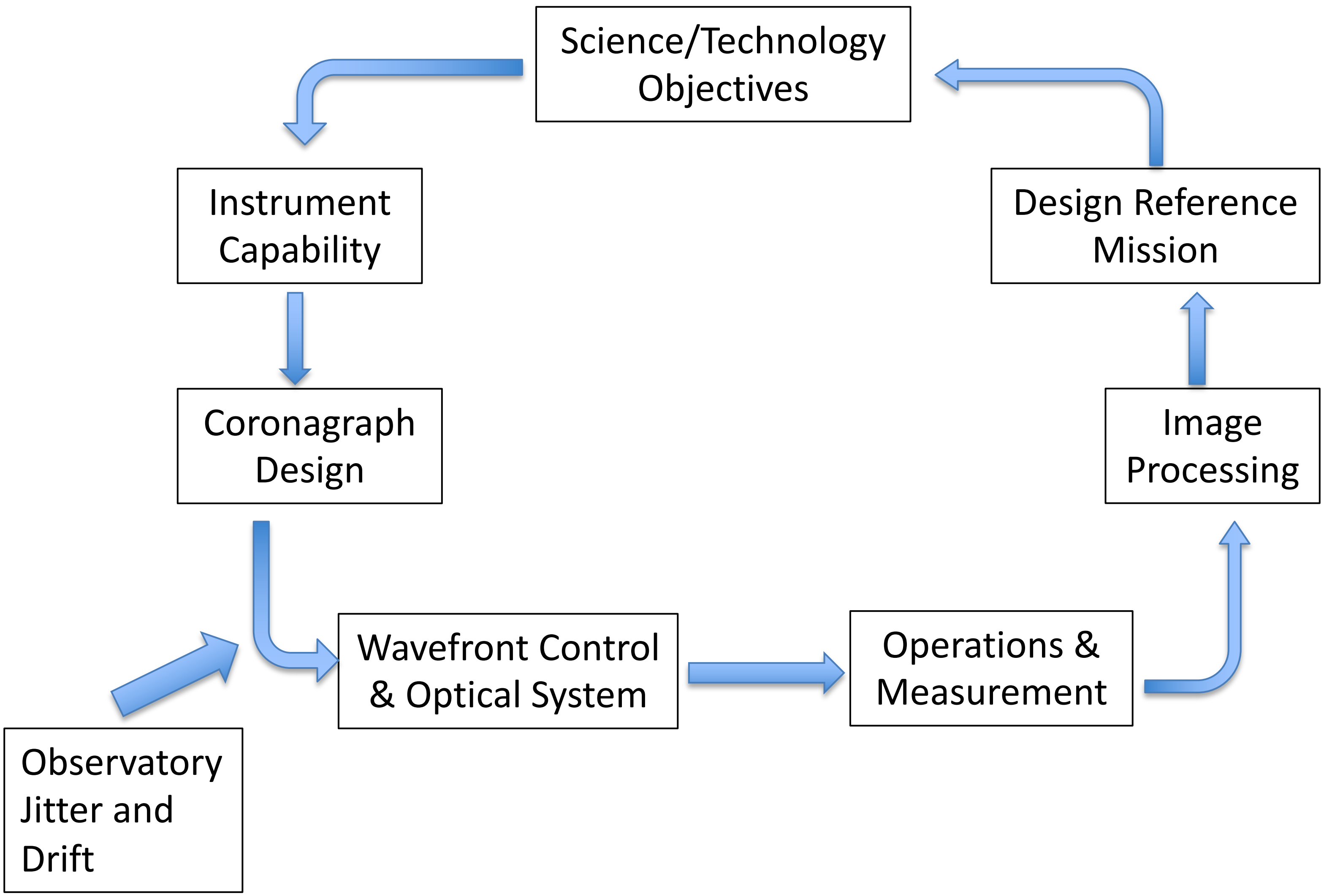
This must be done for angles corresponding to the furthest star at the inner edge of the habitable zone. This is called the Inner Working Angle (IWA).

For Example, for a 6 m telescope imaging a planet at 500 nm and at 60 marcsec, the planet appears at ~3.5 lambda/D relative to the star's PSF (and 1.75 lambda/D at 1000 nm).

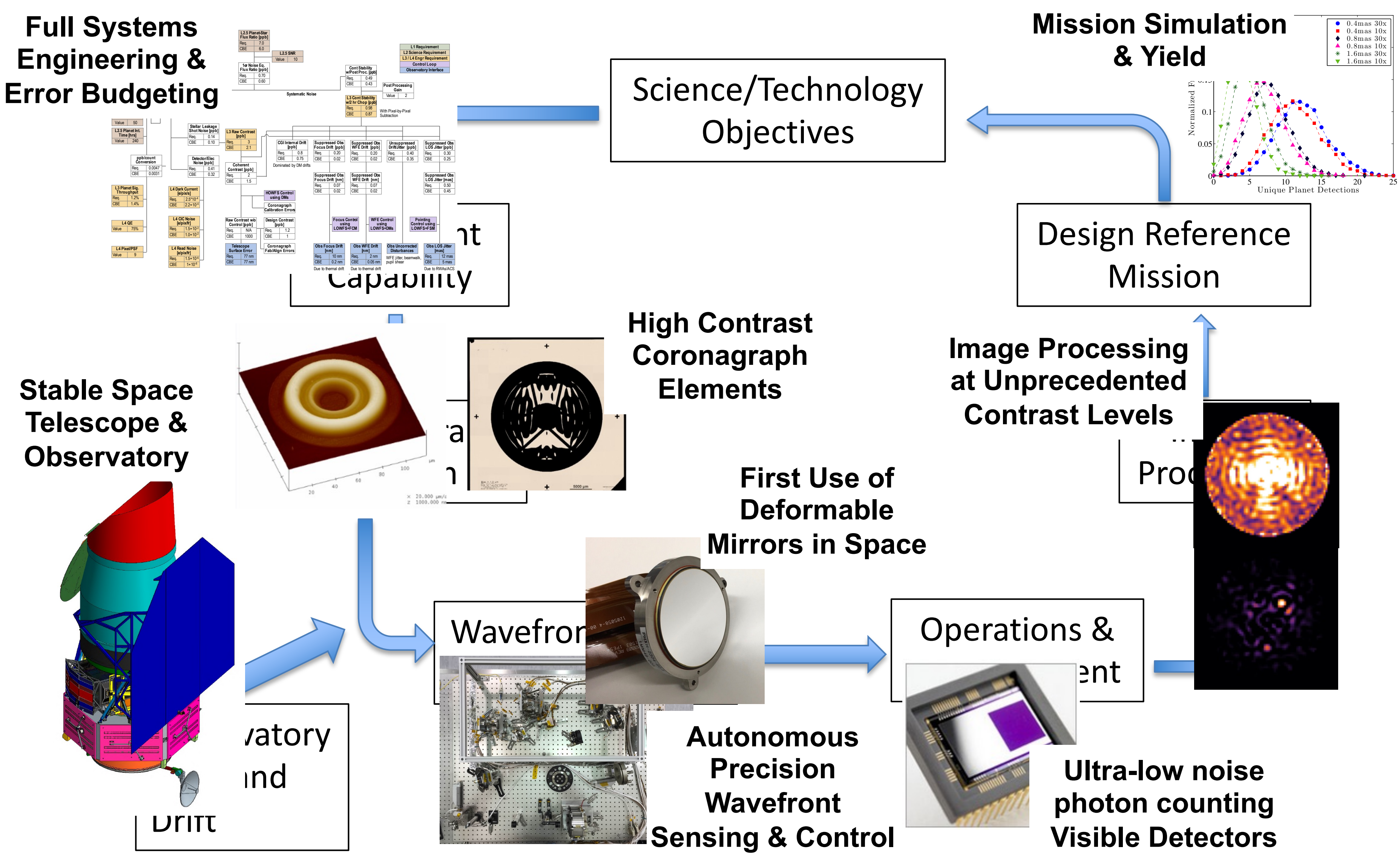
We do this via a coronagraph.



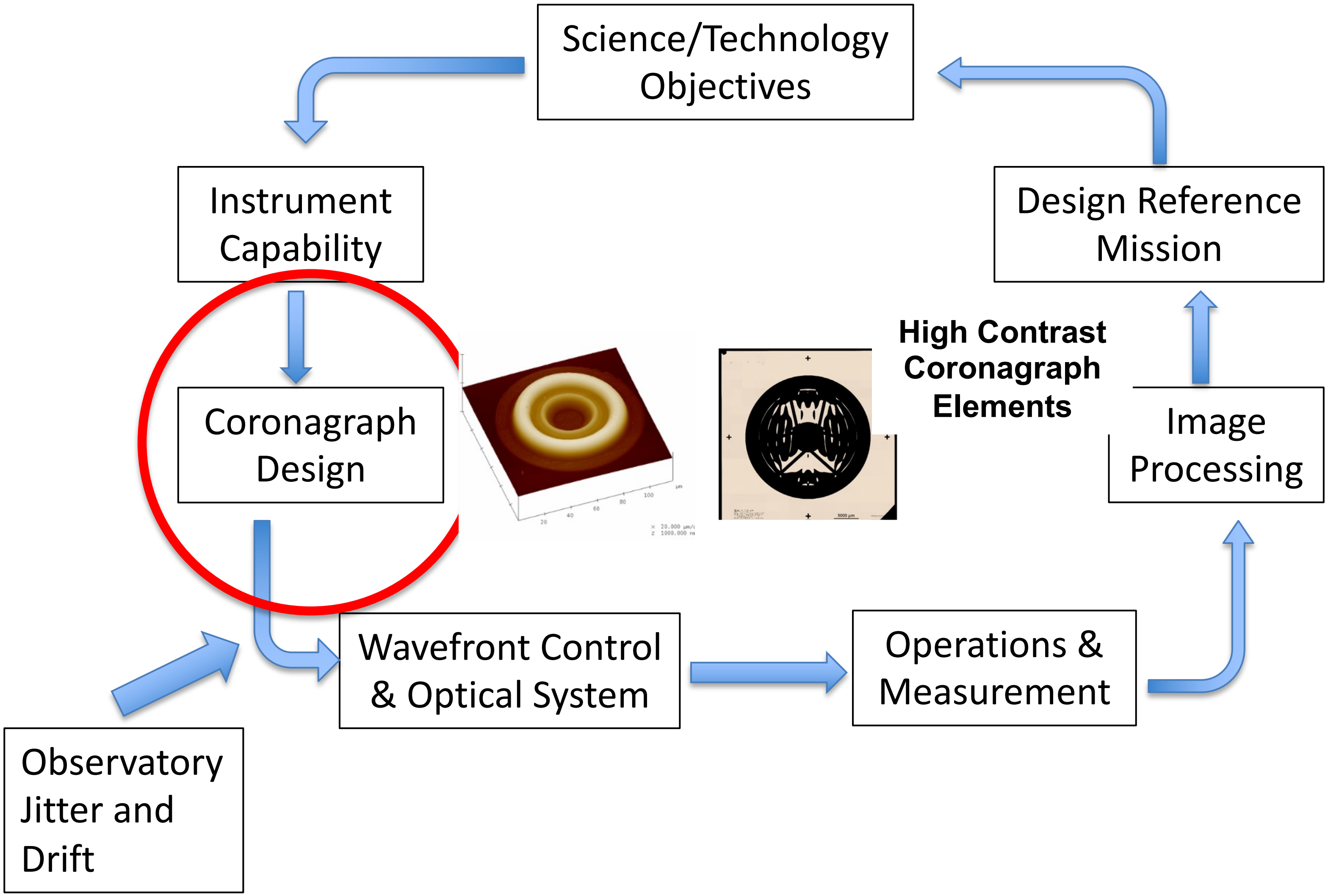
A Coronagraph is a System



The Coronagraph on Roman as a Pathfinder



Coronagraph Design



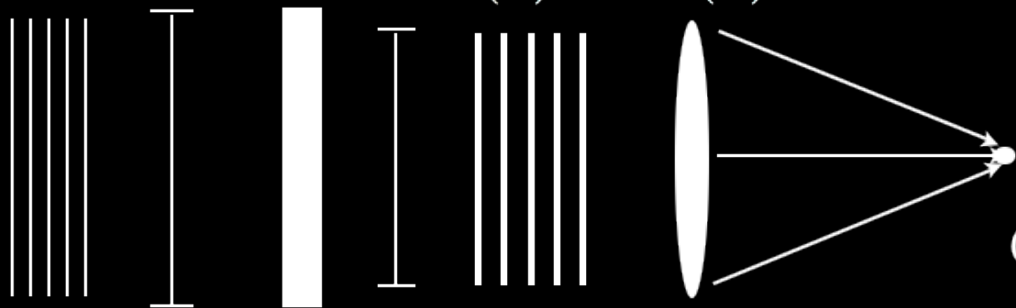
Coronagraph Metrics

- Contrast: The ratio of the peak of the stellar point spread function to the halo at the planet location.
- Inner Working Angle: The smallest angle on the sky at which the needed contrast is achieved and the planet is reduced by no more than 50% relative to other angles.
- Throughput: The ratio of the light in the planet PSF to the nominal telescope PSF after high-contrast is achieved.
- Bandwidth: The wavelengths at which high contrast is achieved.
- Sensitivity: The degree to which contrast is degraded in the presence of aberrations.

Coronagraph performance also differs depending upon aperture (monolith vs. segmented, off-axis vs. on-axis)

Coronagraph Contrast

$E_i \quad \mathcal{A}(x) \quad \mathcal{C}\{E_i\} \quad \mathcal{A}_c(x) = A_c(x)e^{i\psi(x)}$



$$P_c(\omega) = \left|\mathcal{F}\left\{A_c(x)e^{i\psi(x)}\right\}\right|^2$$

On-Axis Point Spread Function

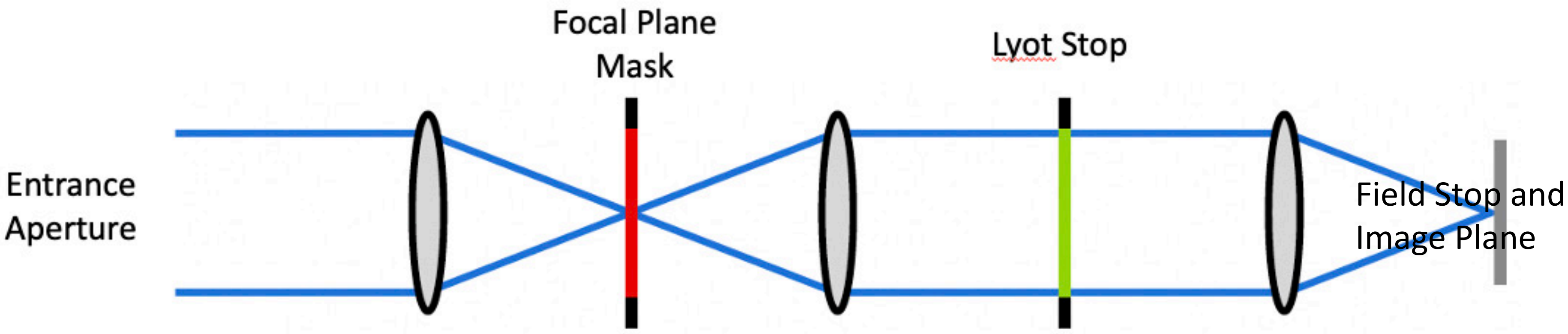
The Instrument Contrast Ratio (at a specific wavelength)

$$C_i = \frac{\int_{\Delta\Omega} P_c(\omega)d\omega}{\Delta\Omega P_o(0)} = \frac{\int_S |\mathcal{A}_c(x)|^2 dx}{\Delta\Omega A_o^2}\left[1 - \frac{\int_{\Delta C} P_c(\omega)d\omega}{\int_{-\infty}^{\infty} P_c(\omega)d\omega}\right]$$

Reduce the exit amplitude

Shift the energy (uncertainty principal)

A Generic Coronagraph

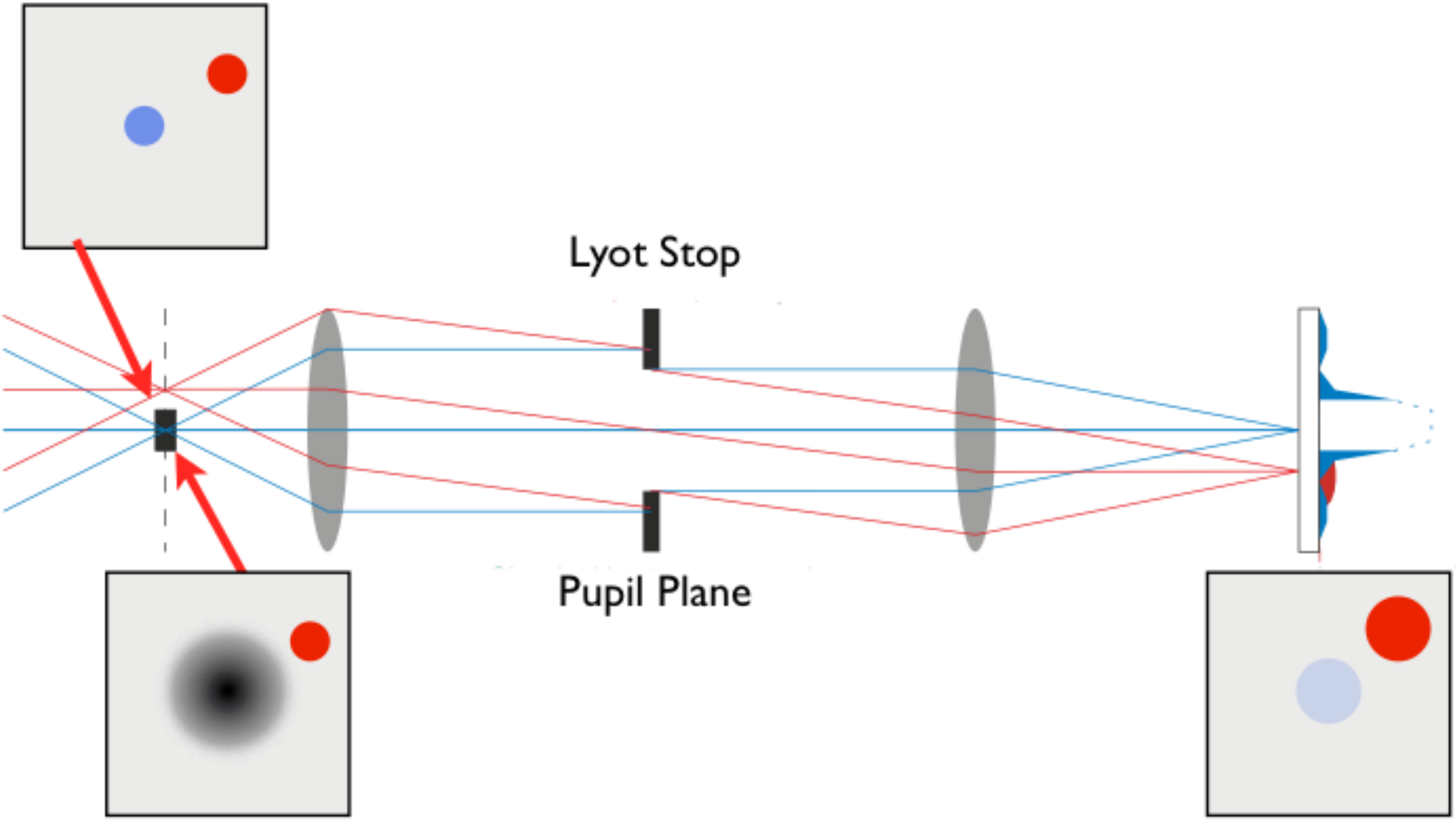


All coronagraphs work by modifying amplitude or phase at the entrance aperture, amplitude or phase at the first focal plane, amplitude or phase at the Lyot plane, or some combination of them.

All are based on using properties of the Fourier Transform.

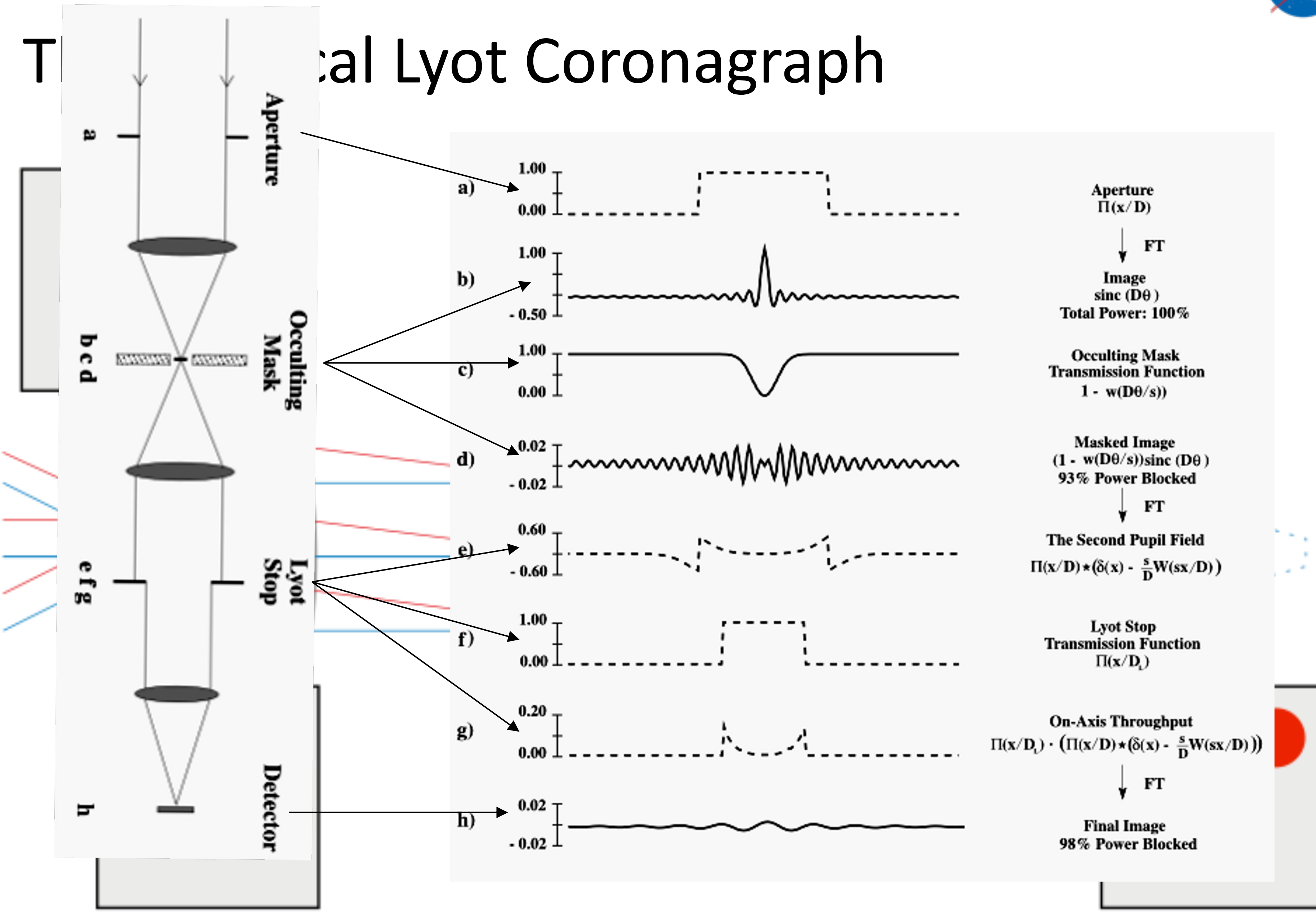
Mennesson et al.

The Classical Lyot Coronagraph



Courtesy Matt Kenworthy,University of Leiden

Tl cal Lyot Coronagraph



Courtesy Matt Kenworthy,University of Leiden

Coronagraph Families

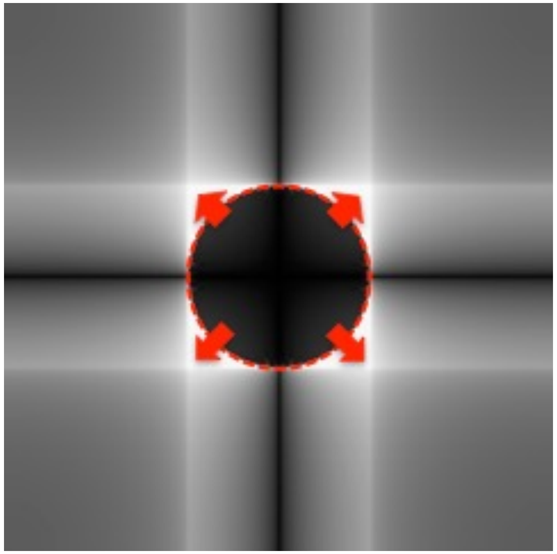
Image Plane

- Lyot & Bandlimited Lyot (Gemini, Keck, Hubble, Subaru, Palomar, VLT, JWST NICI, WFIRST)
- 4 Quadrant Phase Mask (JWST MIRI, VLT, LBT)
- Optical Vortex (Palomar, VLT, LBT), AIC, VNC and other nullers

Pupil Plane

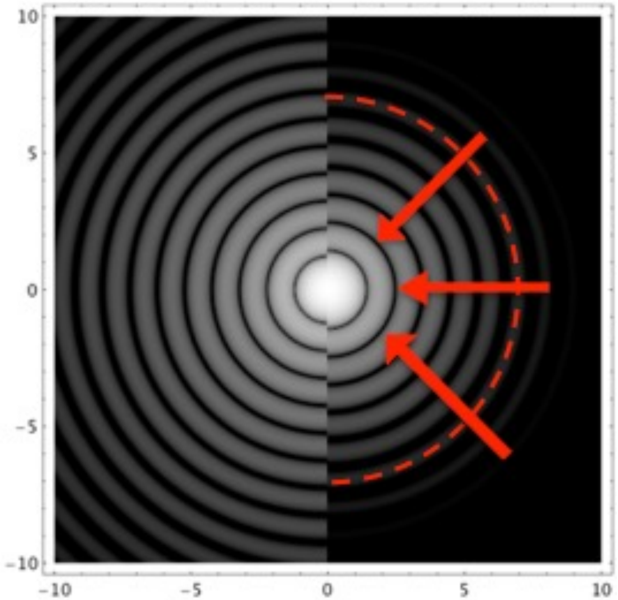
- Apodized pupils(VLT)
- Shaped pupils (Subaru, Roman)
- Pupil remappers (PIAA) (Subaru)
- Apodized phase plate (MMT, Magellan, VLT)

Lyot Plane



APLC, SPLC (GPI, VLT/SPHERE, Palomar)

Image Plane

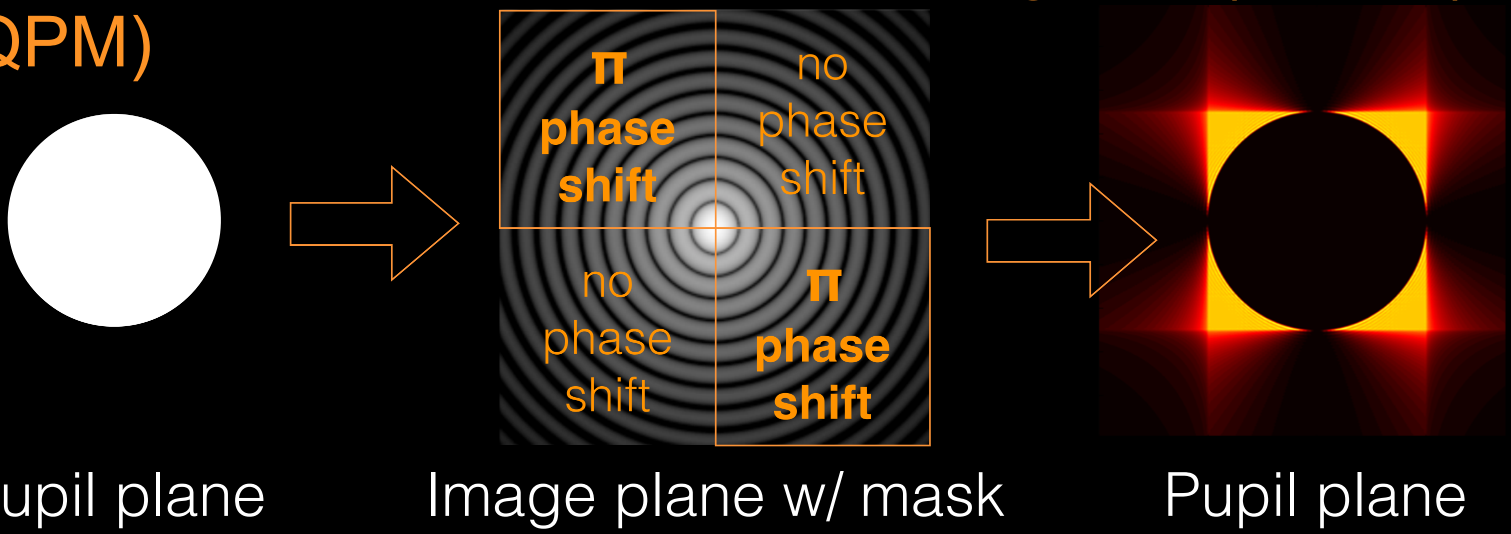


Example Coronagraphs That Change Amplitude

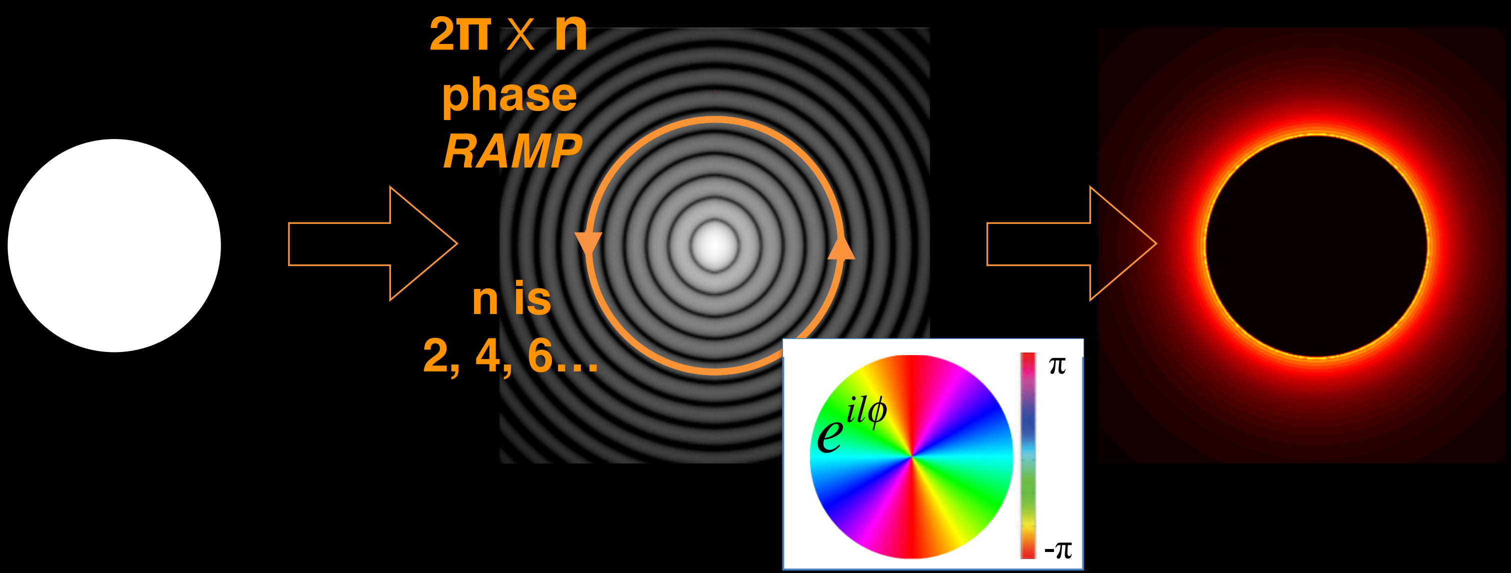
Focal Plane Amplitude Mask: Lyot & Bandlimited Lyot, AIC

Focal Plane Phase Mask: 4QPM, Vector Vortex

Four-Quadrant Phase Mask coronagraph (Rouan) (4QPM)



Vector vortex coronagraph (Mawet)



Example Coronagraphs That Reshape PSF

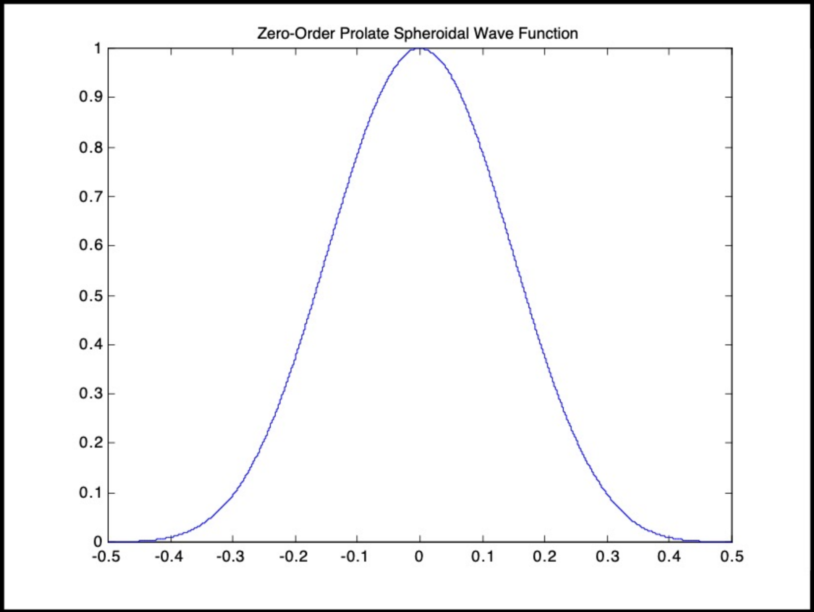
Pupil Plane Amplitude Mask: Apodization, Shaped Pupils, PIAA

Pupil Plane Phase Mask: APP

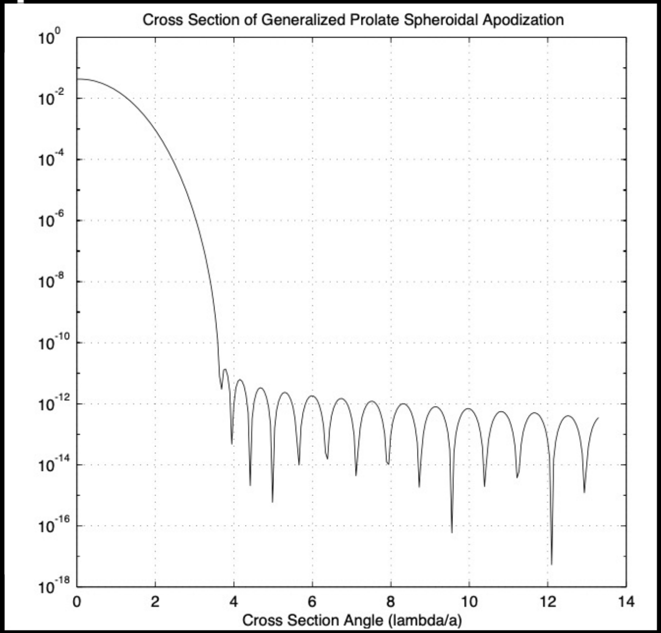
Pupil Apodization to Reshape PSF

Slepian, D., “Analytic Solution of Two Apodization Problems”, September, 1965

Pupil Apodization

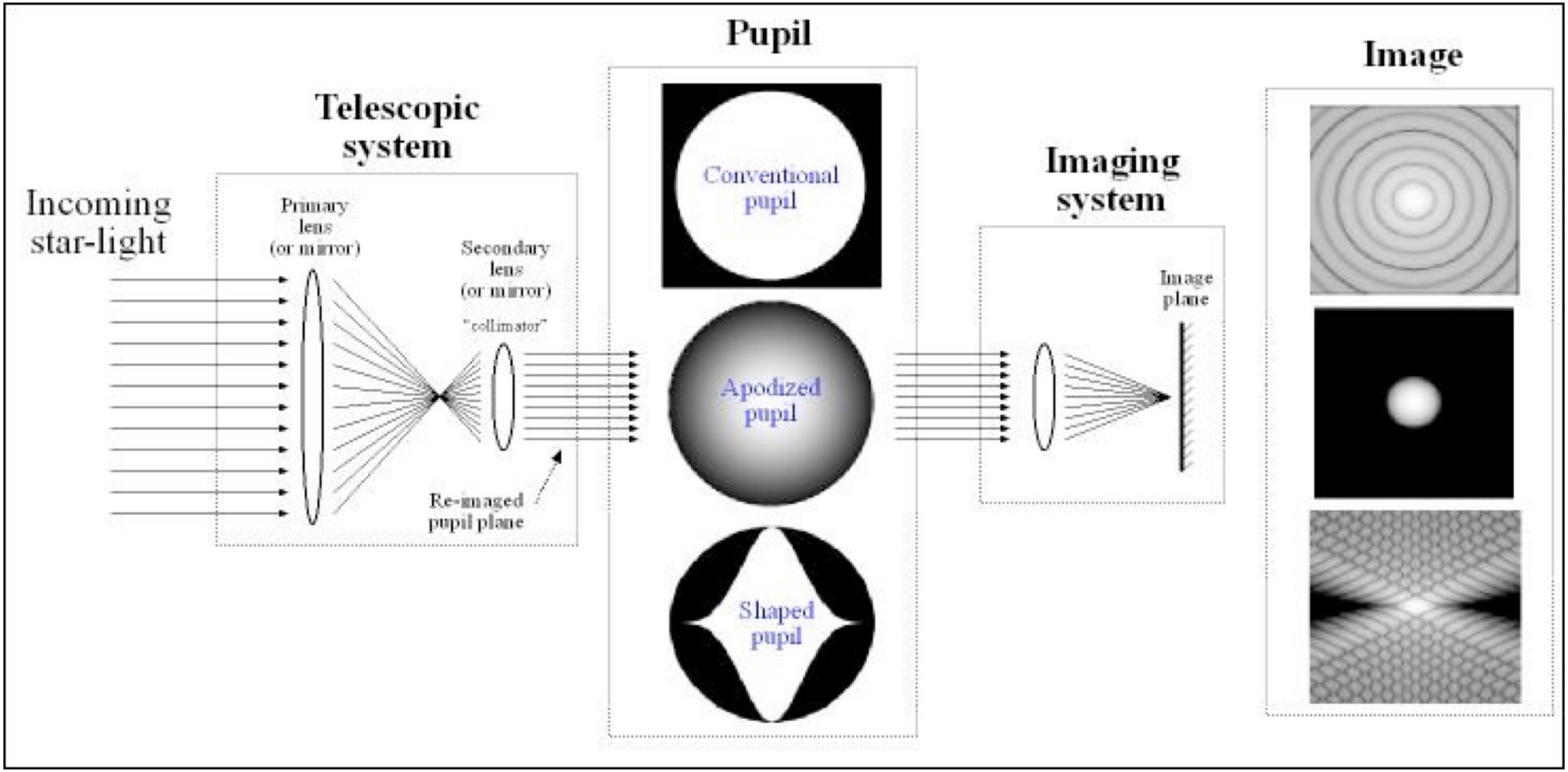


Point Spread Function

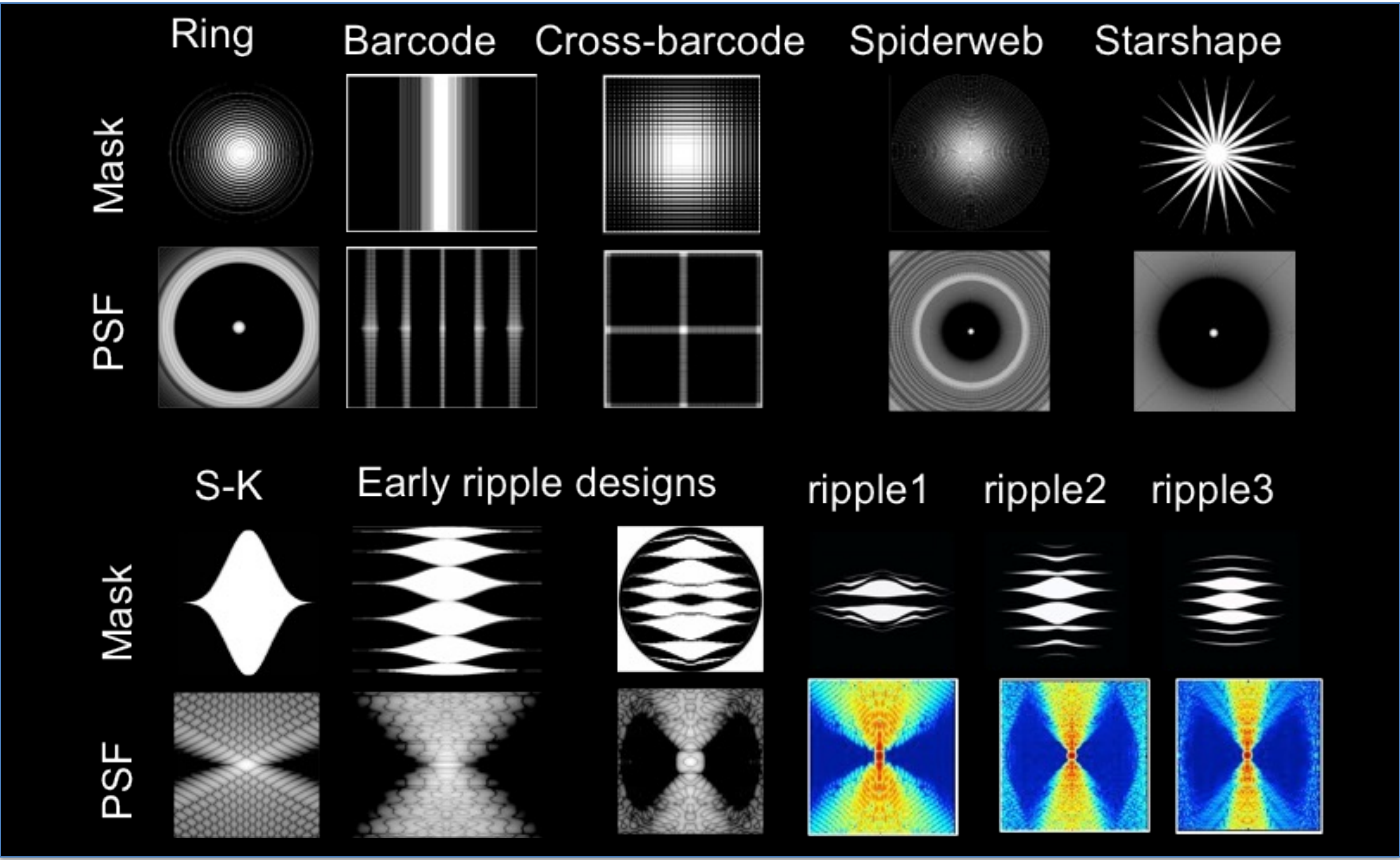


The “optimal” apodization that maximally concentrates light is the Prolate Spheroidal Wavefunction, based on finite uncertainty principle.

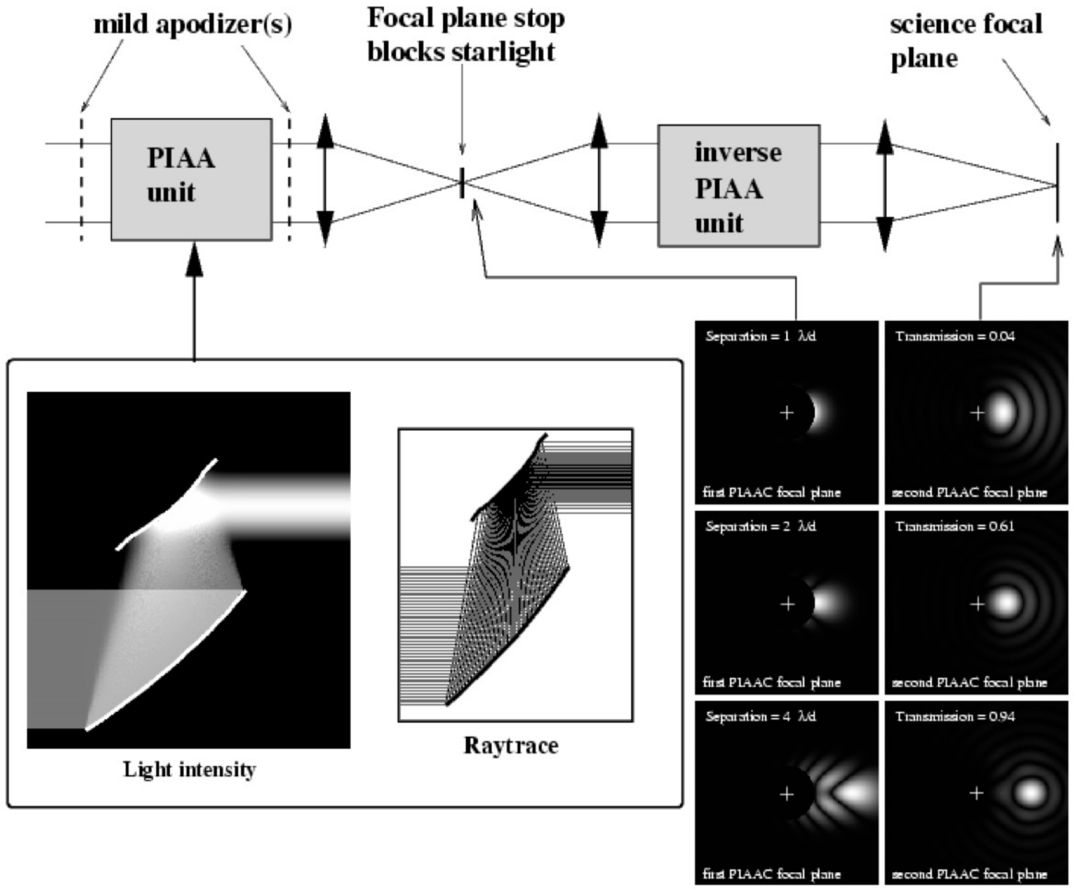
A Shaped Pupil



Shaped Pupils



Pupil Mapping (PIAA)



Pupil Mapping for Apodization

Nearly 100% throughput
100% search area
small (<2 lambda/d) Inner Working Angle

Guyon (2003), Vanderbei & Traub (2003, 2005)

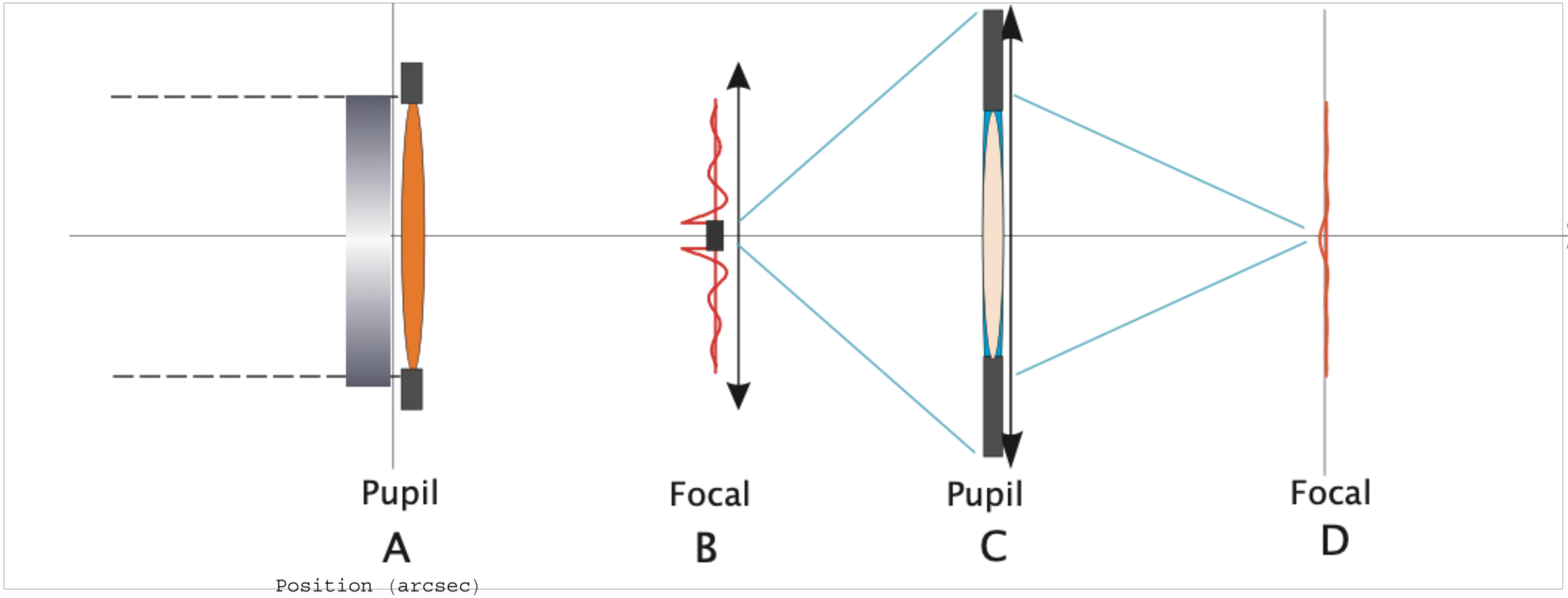
Example Coronagraphs that do a bit of both

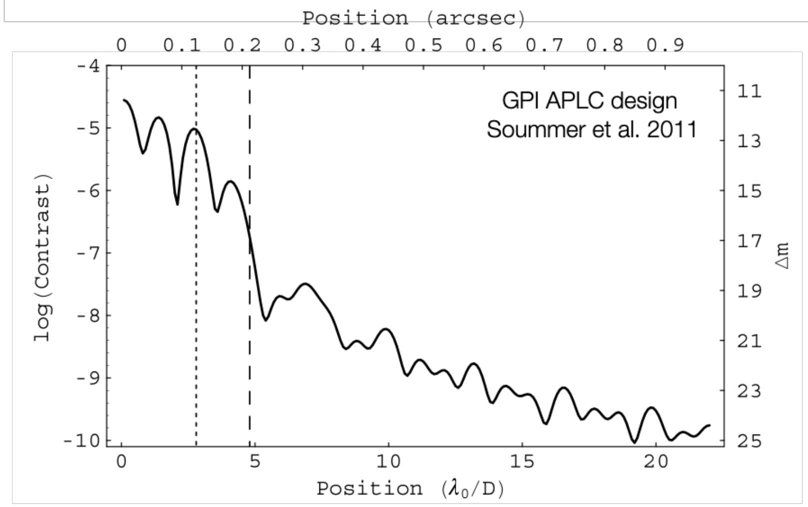
Apodized Pupil Lyot Coronagraph (APLC)

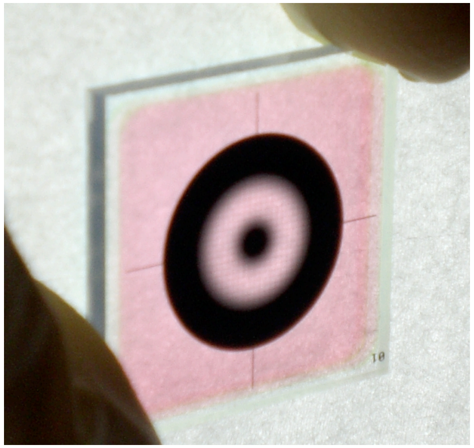
Shaped Pupil Lyot Coronagraph (SPLC)

Apodized Pupil Lyot Coronagraph

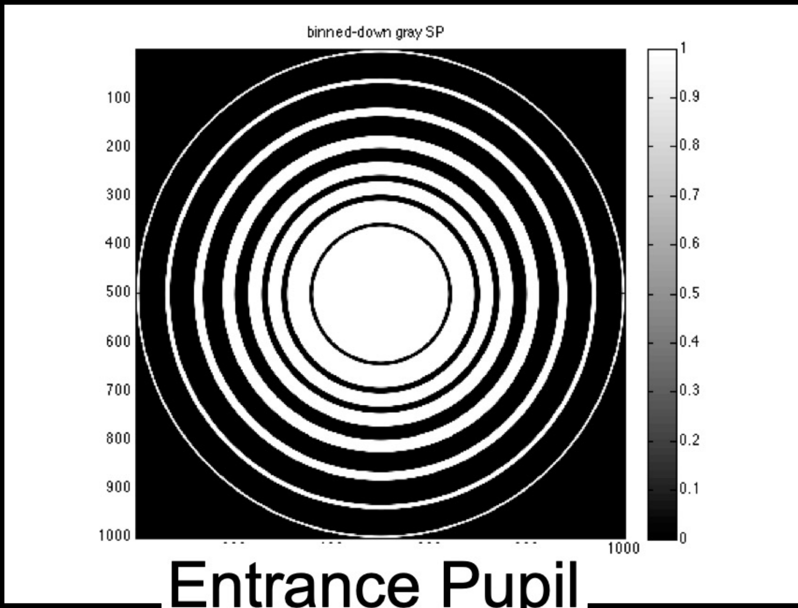
Soummer et al. 2005, 2009, 2011



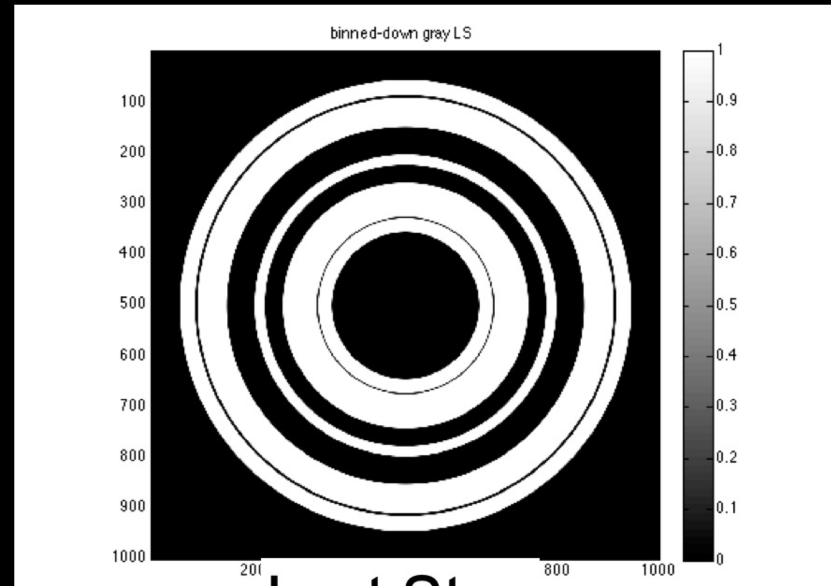




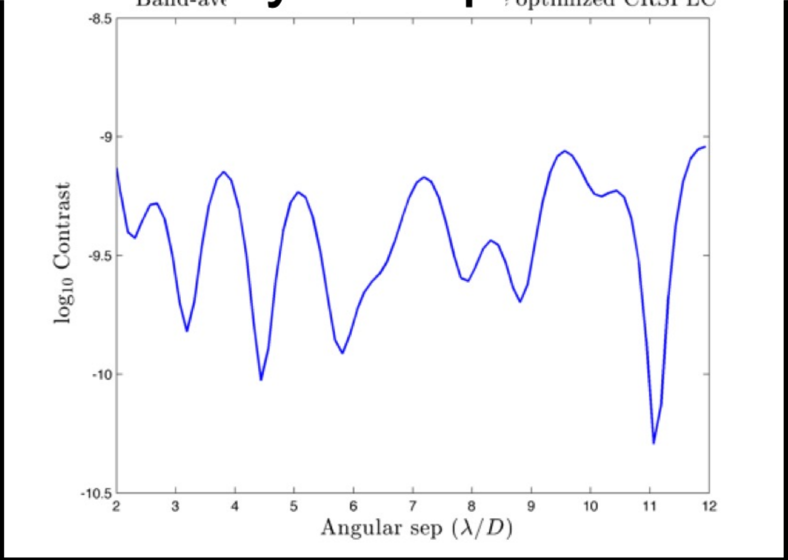
Shaped Pupil Lyot Coronagraph



Entrance Pupil



Lyot Stop



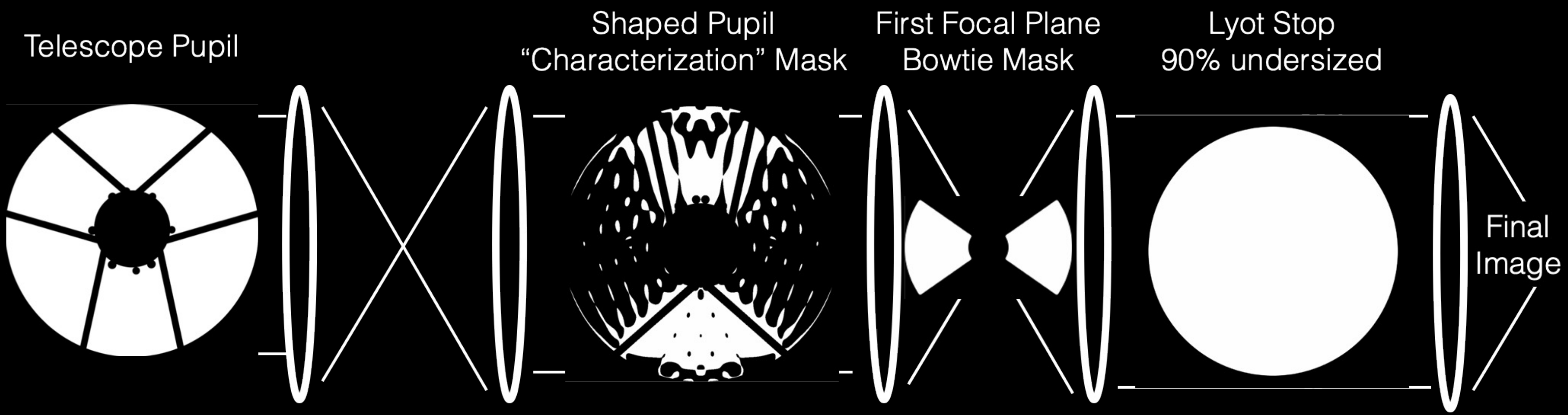
Simultaneously optimize pupil and Lyot plane

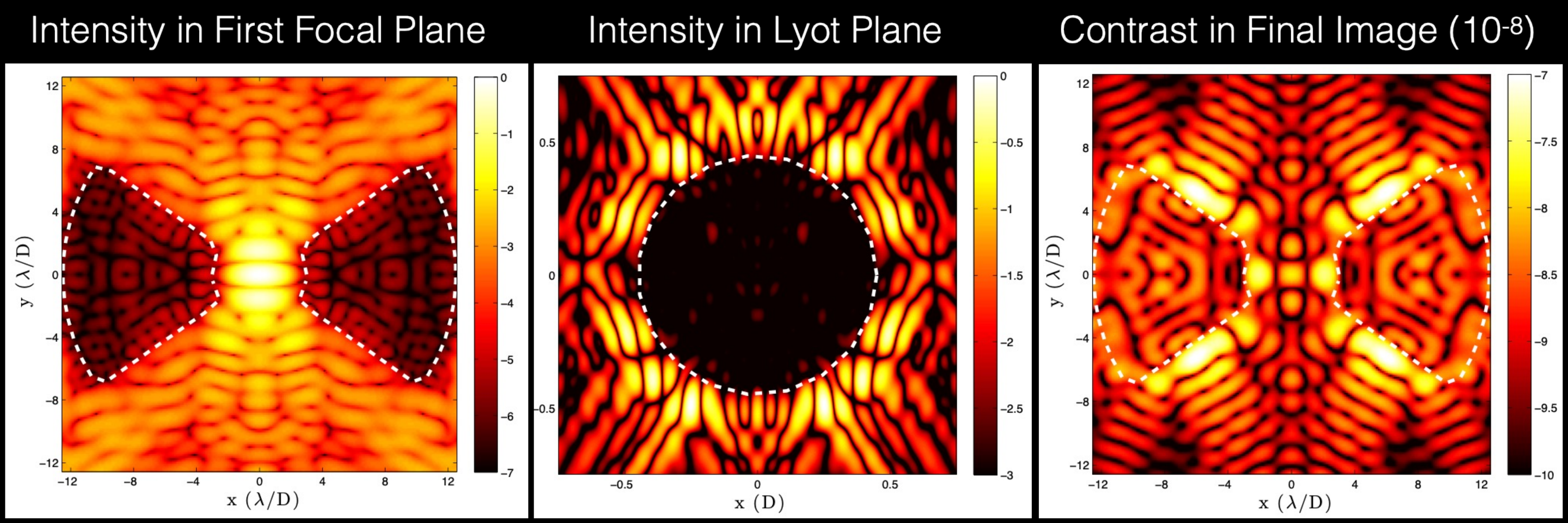
Gains smaller iwa and more throughput

from Neil Zimmerman

Shaped Pupil Lyot Coronagraph for Roman

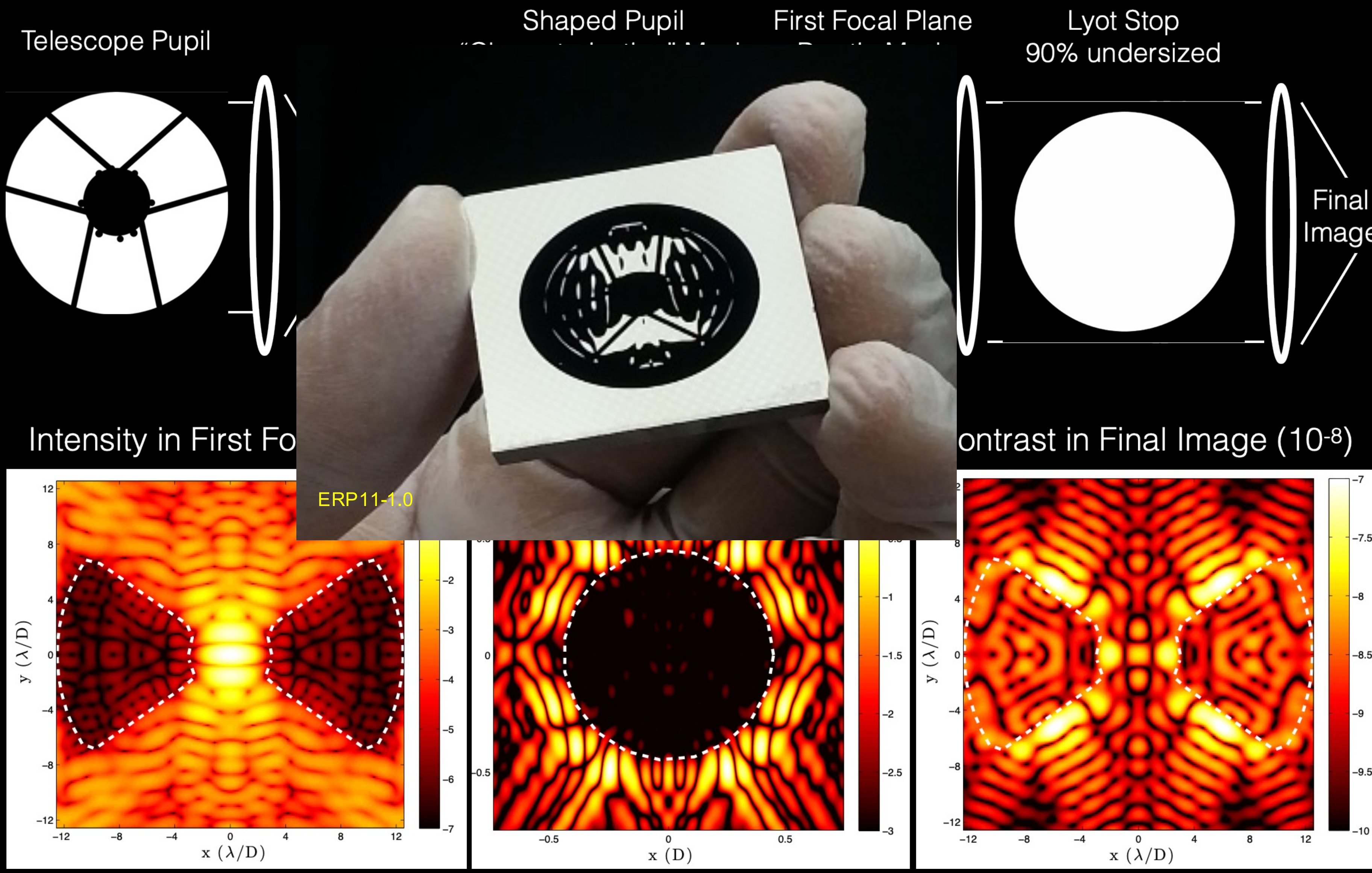
From Neil Zimmerman





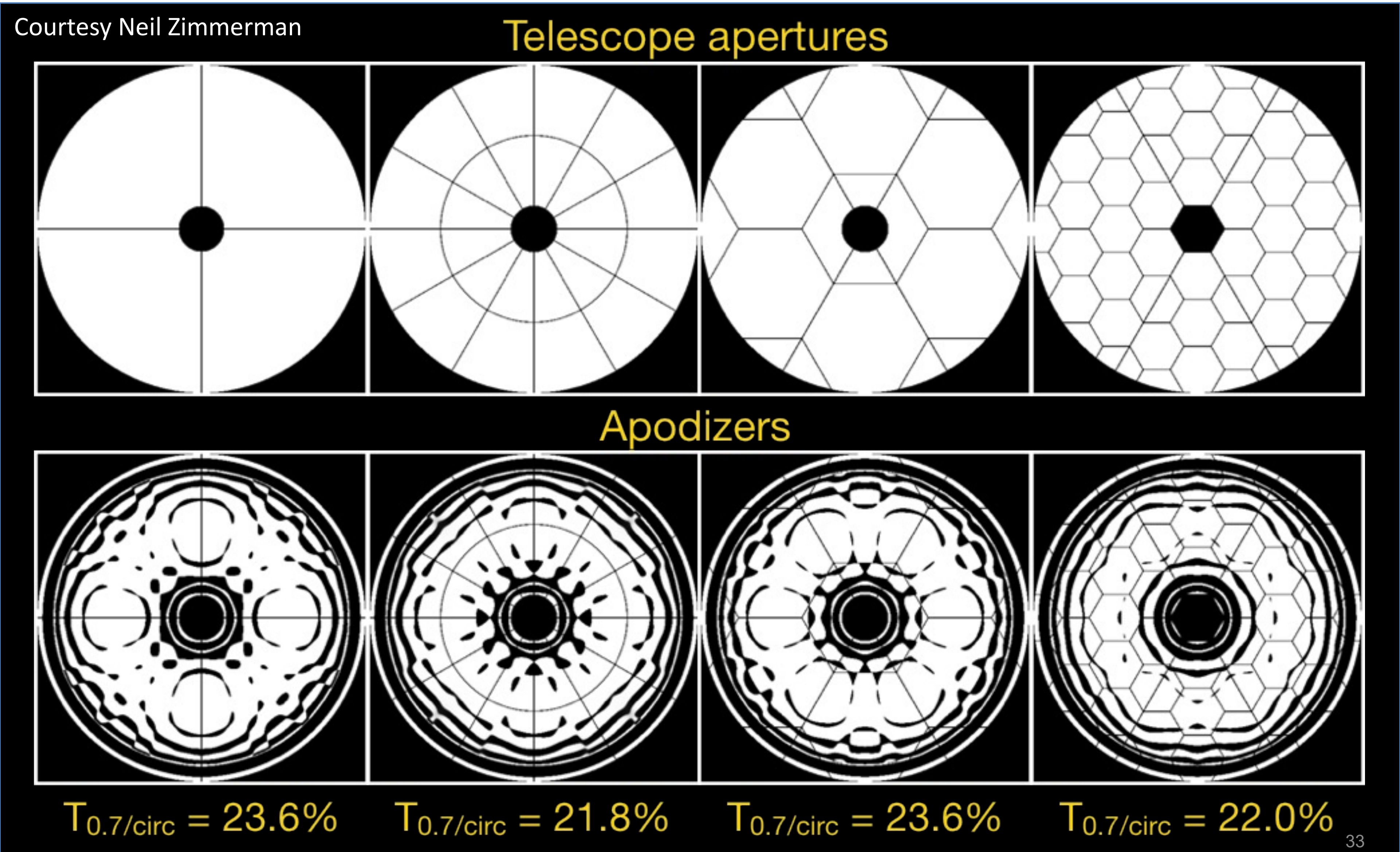
Shaped Pupil Lyot Coronagraph for Roman

From Neil Zimmerman



Segmented and Obstructed Pupils
Apodized/Shaped Pupil Lyot Coronagraph

STScI



Coronagraph Metrics

- Contrast: The ratio of the peak of the stellar point spread function to the halo at the planet location.
- Inner Working Angle: The smallest angle on the sky at which the needed contrast is achieved and the planet is reduced by no more than 50% relative to other angles.
- Throughput: The ratio of the light in the planet PSF to the nominal telescope PSF after high-contrast is achieved.
- Bandwidth: The wavelengths at which high contrast is achieved.
- Sensitivity: The degree to which contrast is degraded in the presence of aberrations.

Coronagraph performance also differs depending upon aperture (monolith vs. segmented, off-axis vs. on-axis)

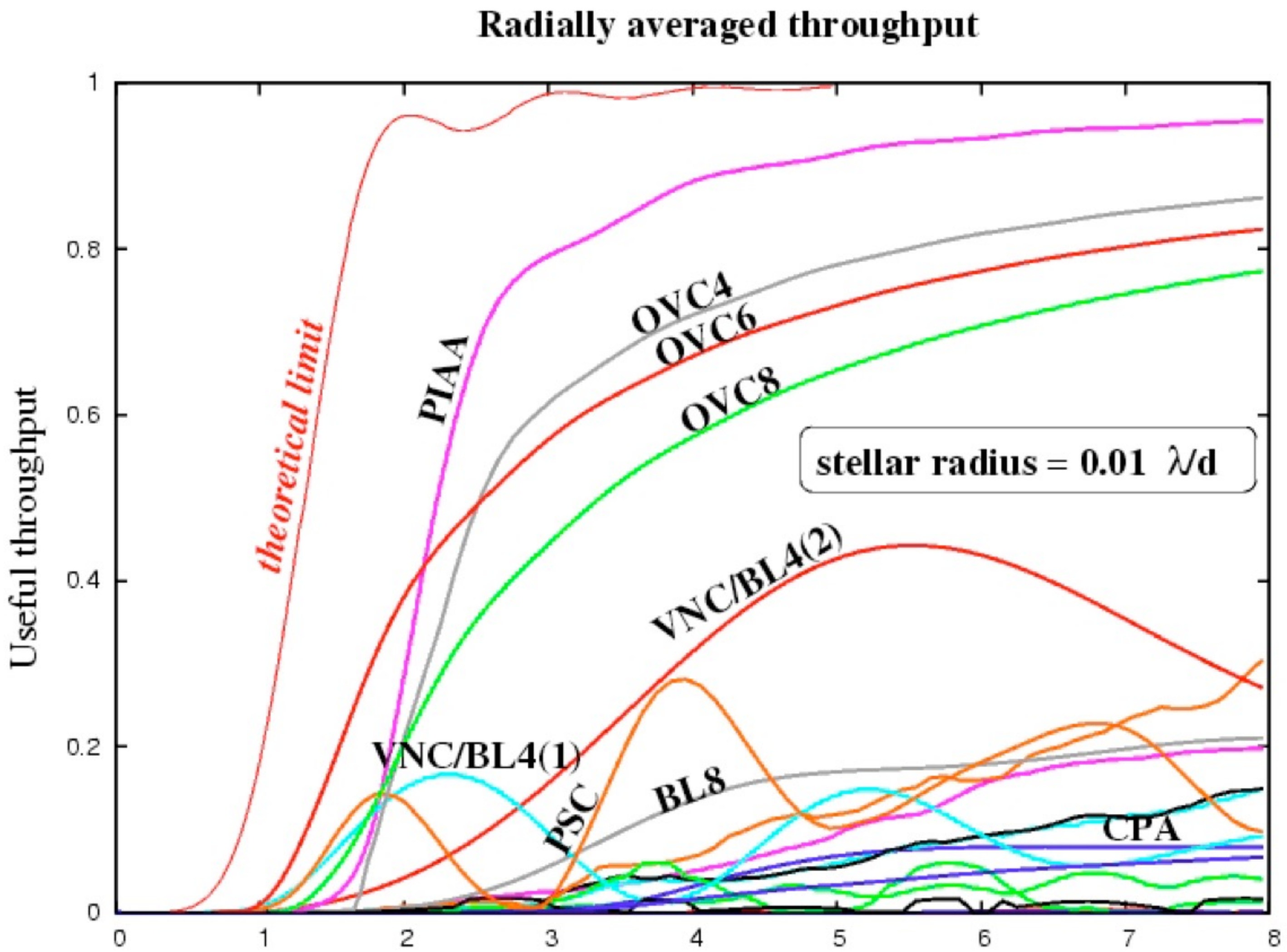
Throughput Definitions

Total Throughput: The ratio of the total planet light in the image plane to the total amount of light without a coronagraph.

Core Throughput: The ratio of the light in the central core of the planet PSF to the total amount of light without a coronagraph.

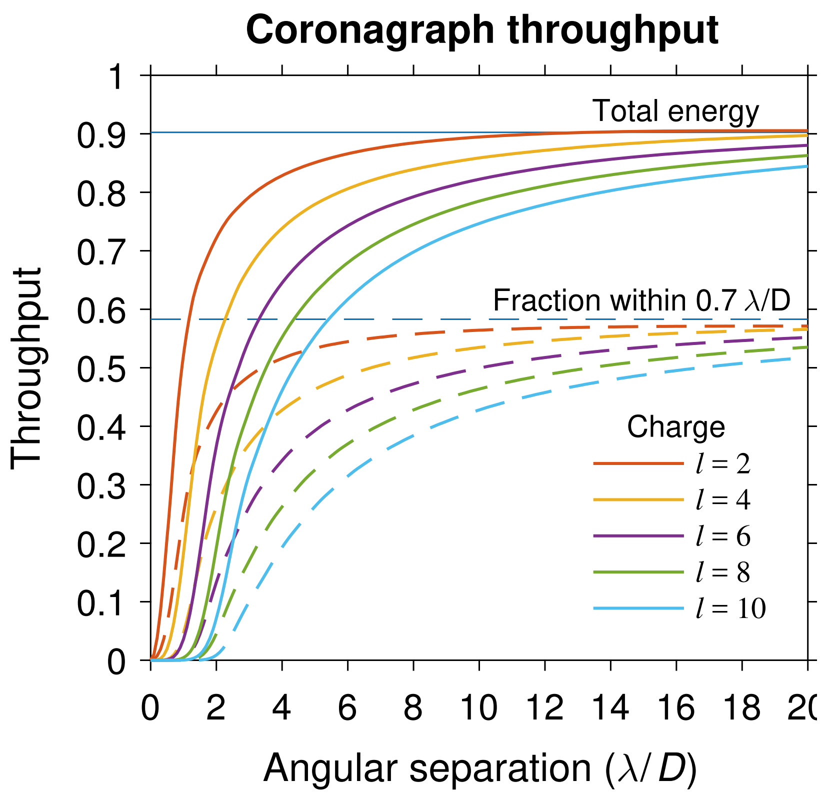
Useful Throughput (Guyon et al. 2006): The maximum fraction of planet light that can be separated from starlight.

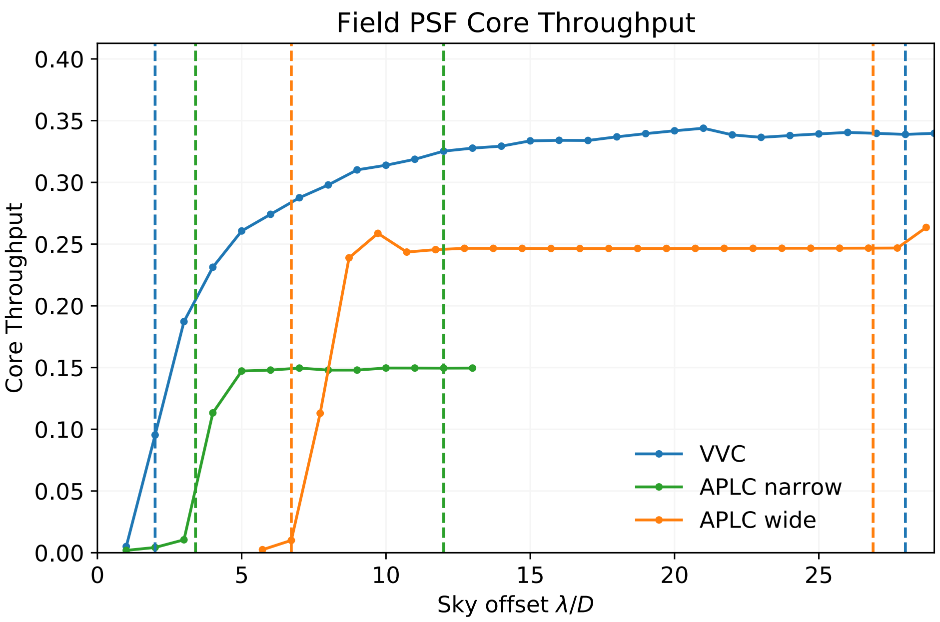
Throughput and Inner Working Angle



Guyon, Pluzhnik, Kuchner, Collins & Ridgway
2006, ApJS 167, 81

Ex.: VVC and APLC Throughput





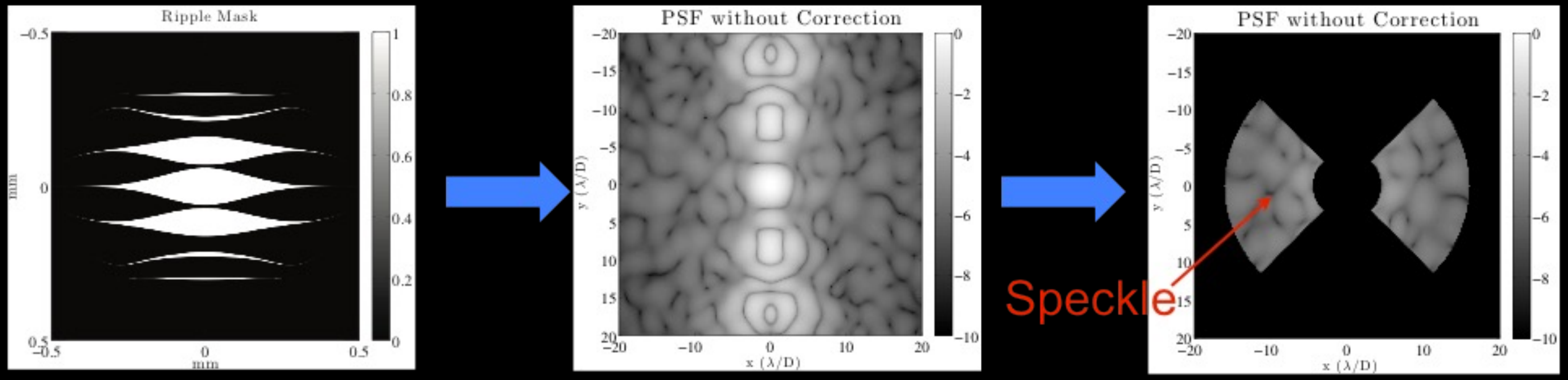
Ruane et al.

Juanola-Parramon et al.

But . . .

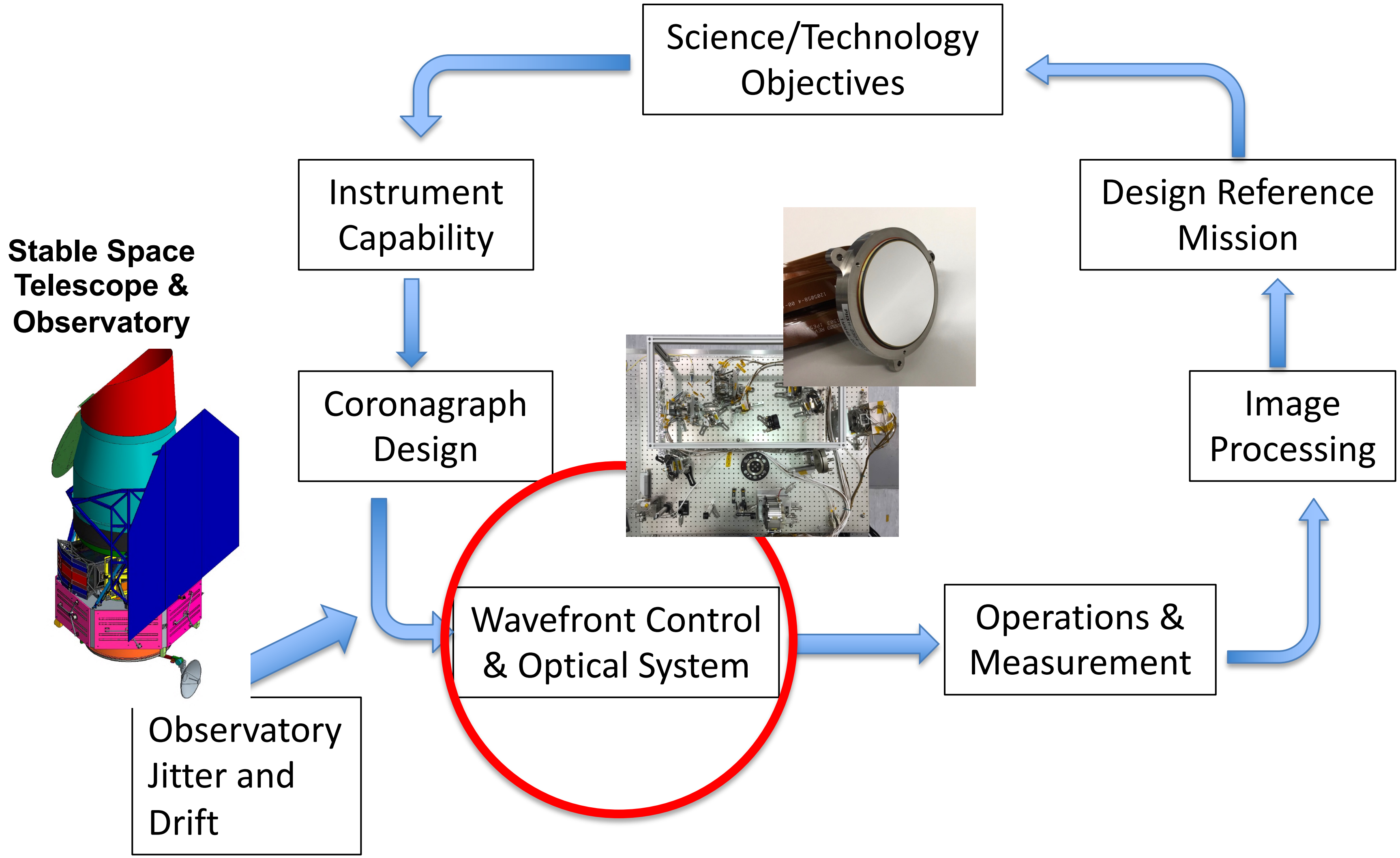
Wavefront Aberrations

Atmospheric distortions and imperfect optics degrade contrast

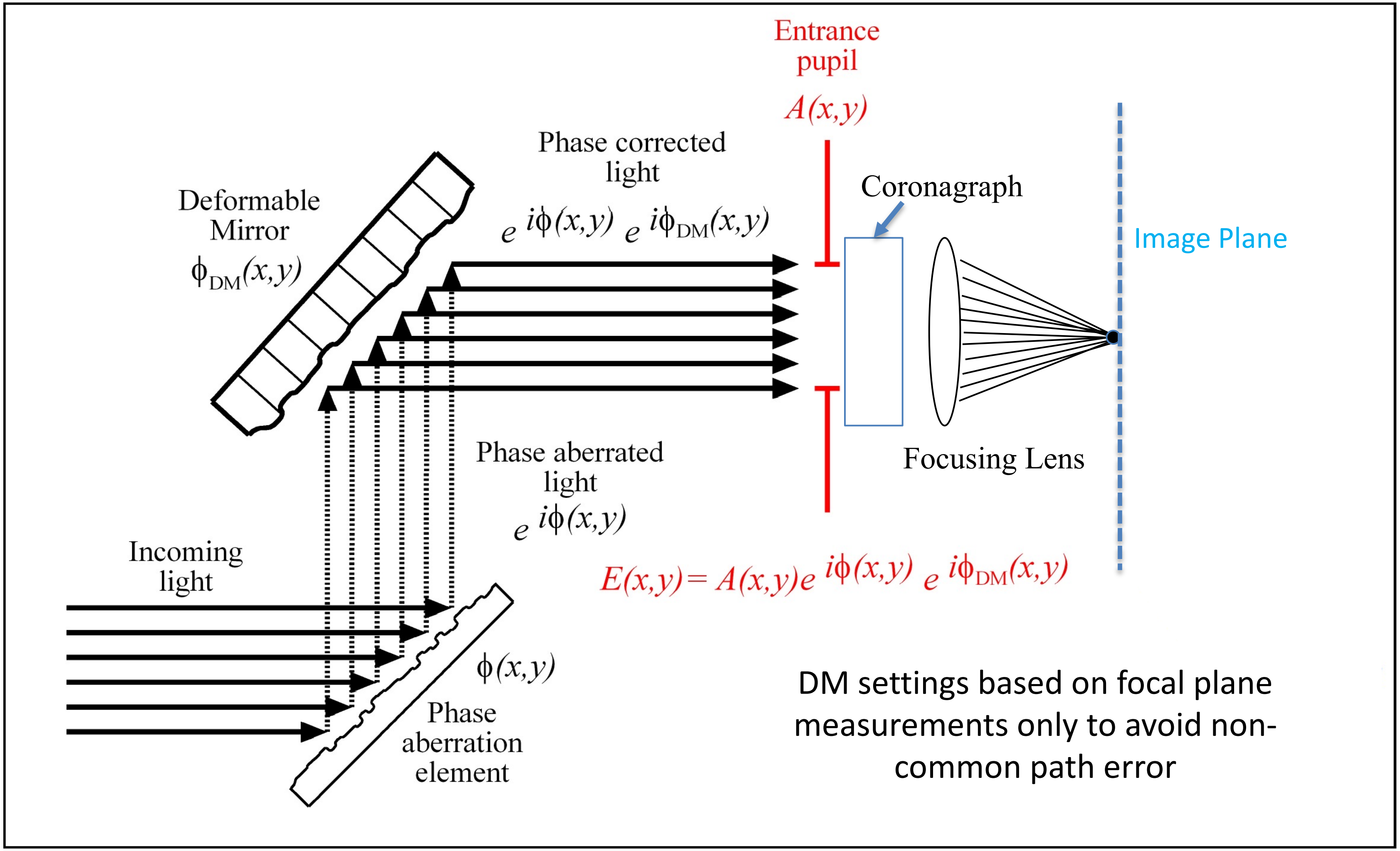


Aberrations significantly degrade contrast: 10^{10} ~10^{5}

Wavefront Estimation and Control



Correcting Phase Only Aberrations



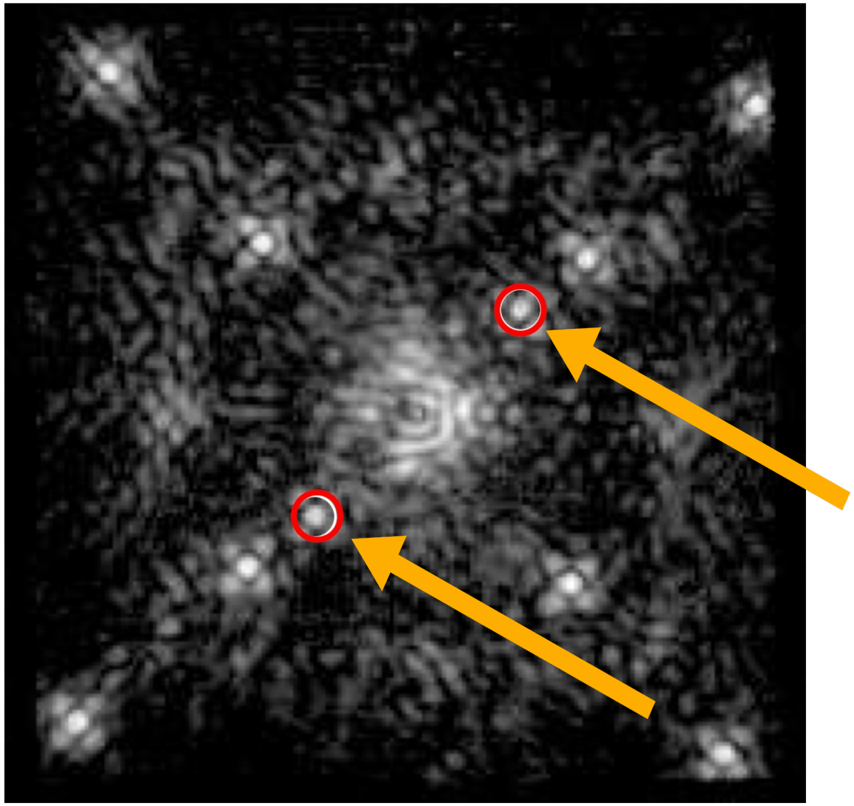
Note: small displacements result in large phase errors in UV and small errors in NIR, setting need for high stroke and high resolution DMs.

Example – Speckle Nulling

A sinewave on a deformable mirror produces a pair of spots in the image plane => we can back out the phase error (DM shape needed) from the science camera image intensity pattern

Deformable mirror

PSF

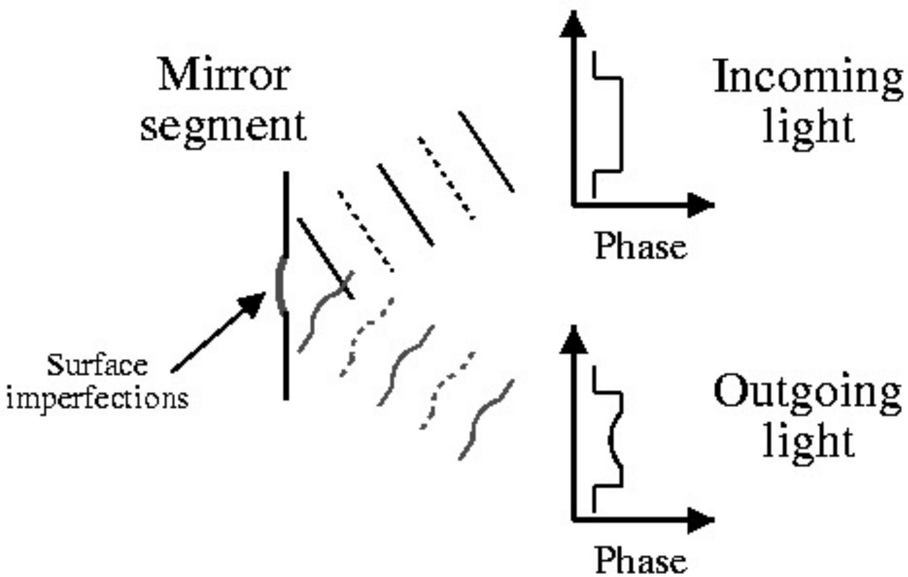


The Fourier transform of a sinewave is a delta function

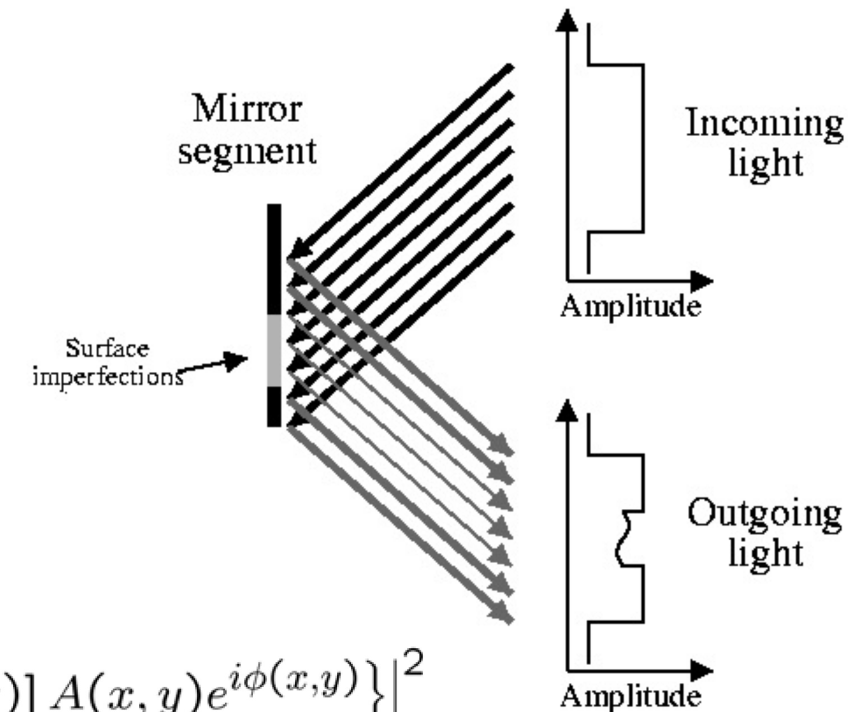
Courtesy of Vanessa Bailey

Phase and Amplitude Errors

Phase Aberrations



Amplitude Aberrations



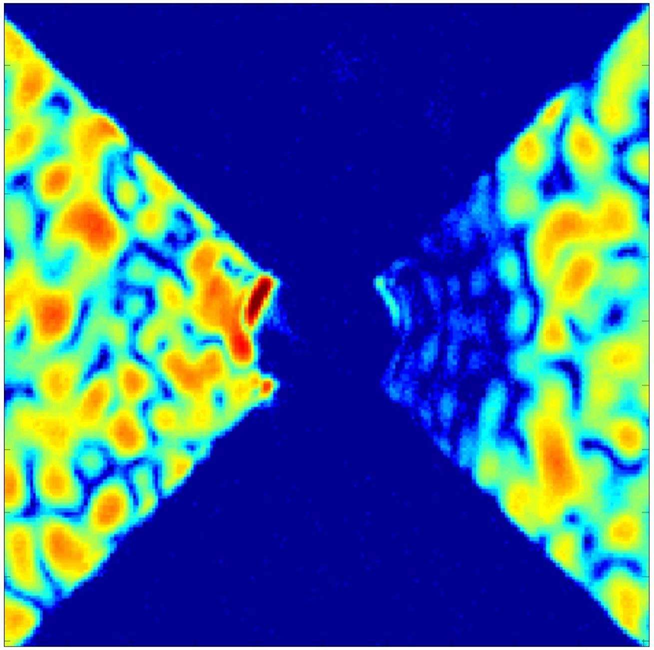
$$I(\xi,\eta) = \left| \mathcal{F}\left\{ [1-\beta(x,y)]\, A(x,y) e^{i\phi(x,y)} \right\} \right|^2$$

Amplitude aberration — Phase aberration

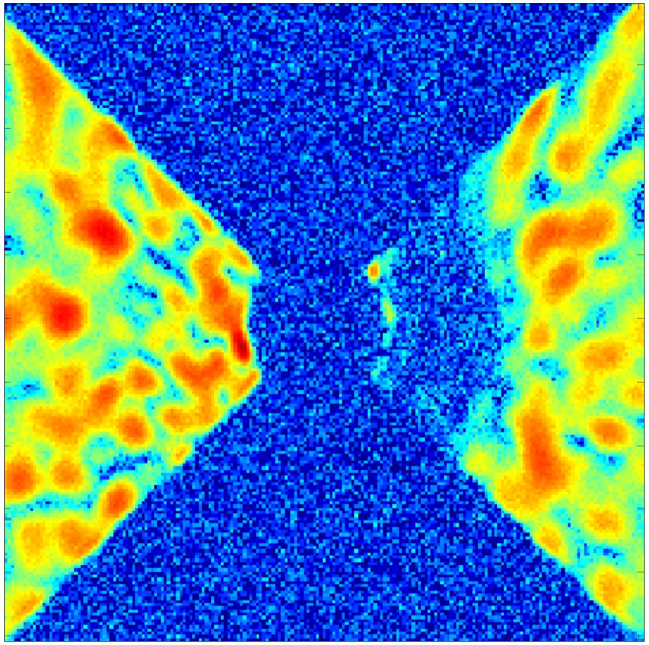
Amplitude errors cannot be corrected with a single DM.

Example: One Sided Dark Hole correcting Amplitude and Phase

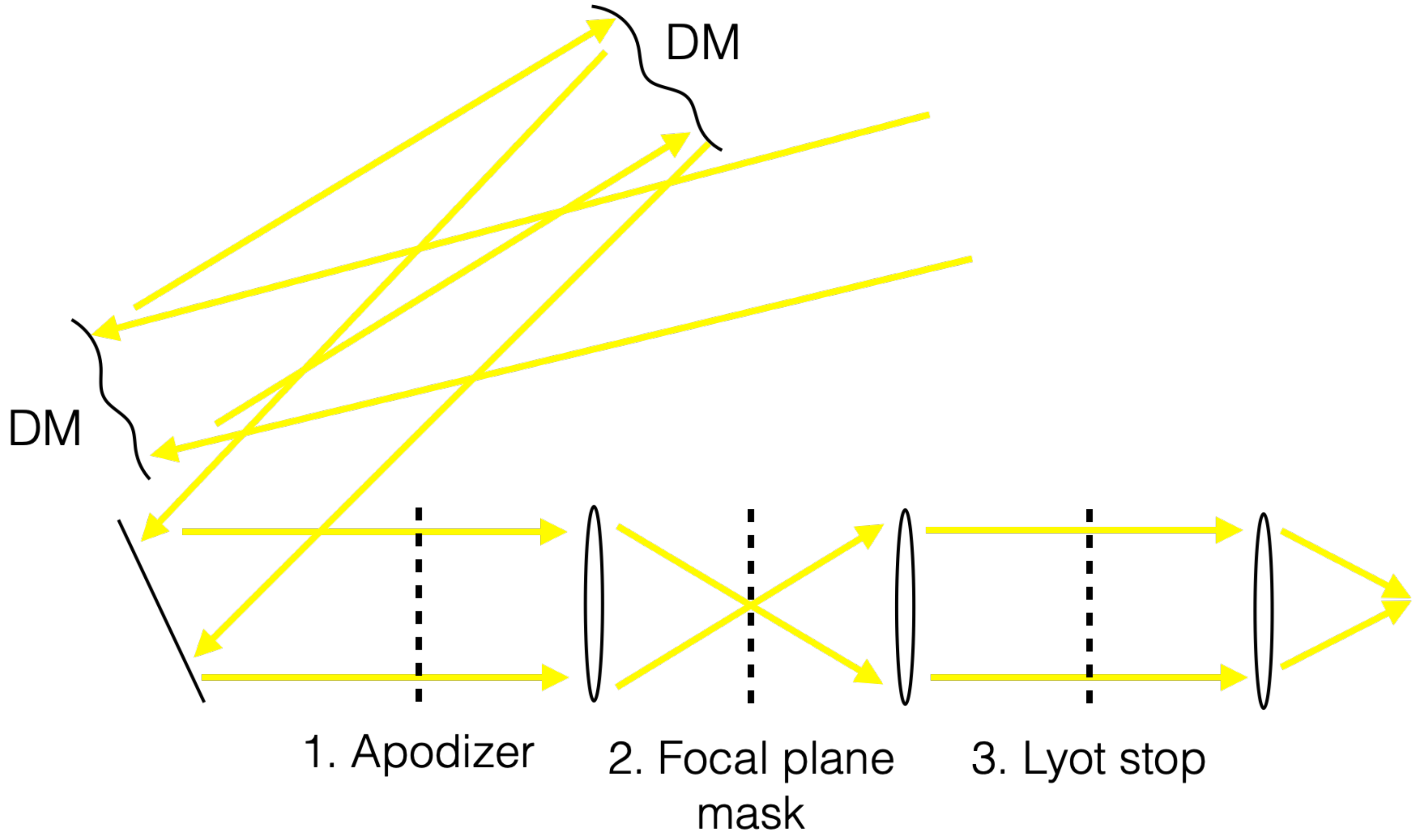
836nm



760nm-840nm

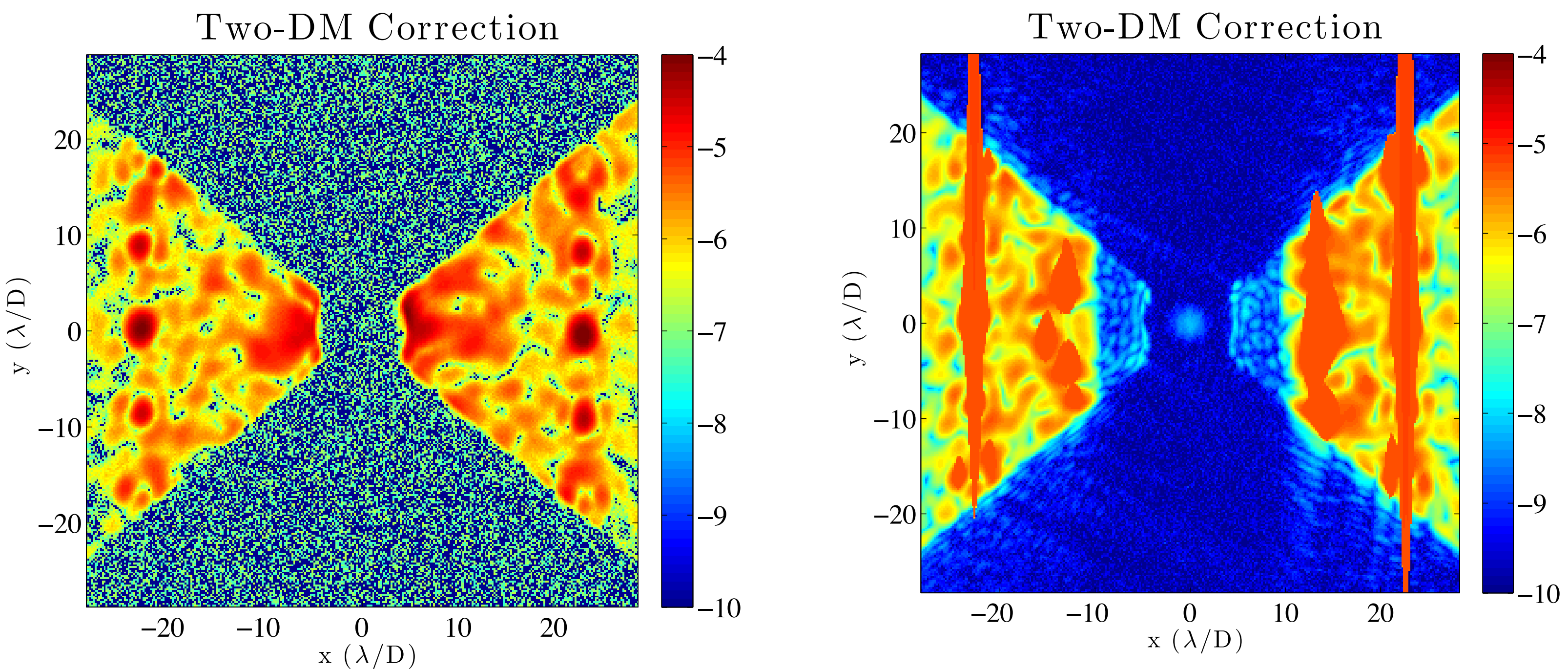


2 DMs for Full Dark Hole (two-sided)



2 DMs for Full Dark Hole (two-sided)

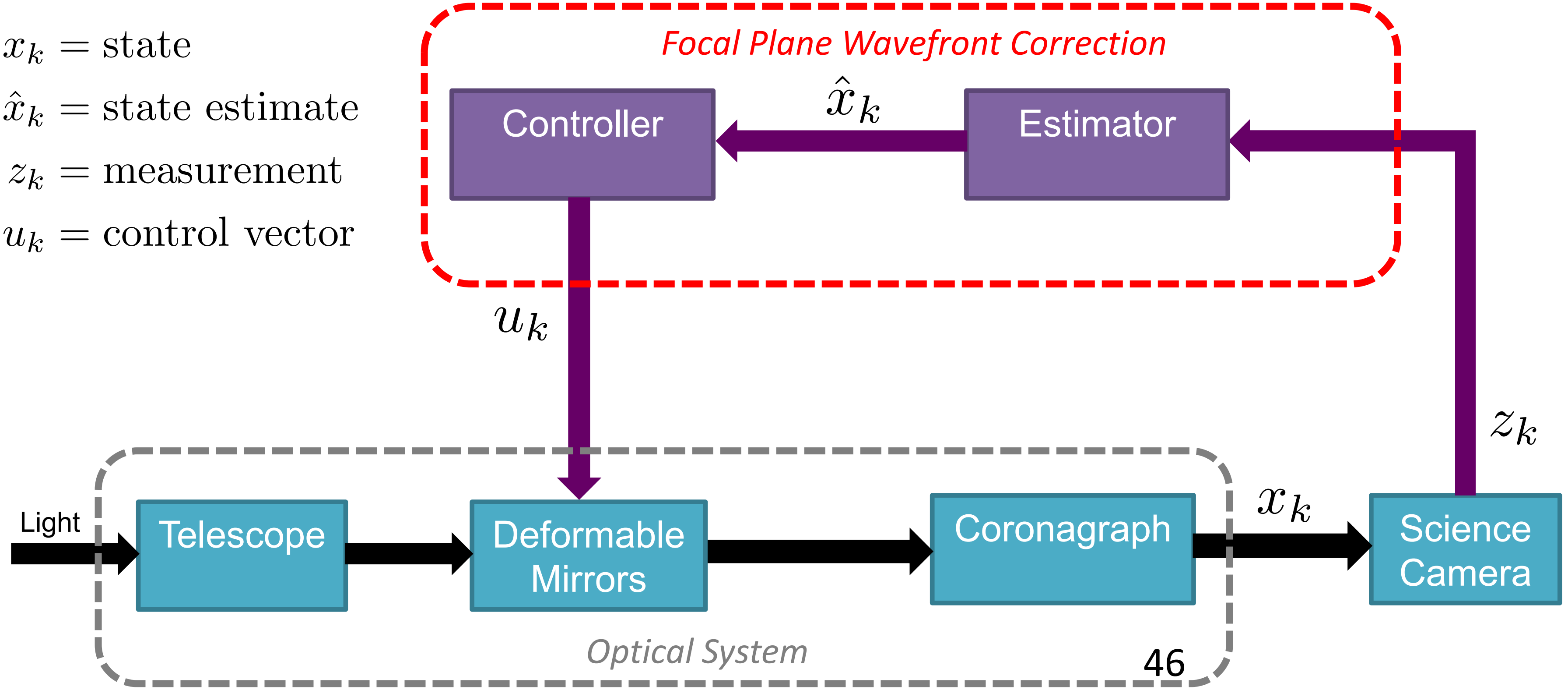
- With 2 DMs in series, both sides of image correctable in phase and amplitude [Shaklan and Green, 2006; Pueyo, Kay, et al. 2009]
- Choose DM separation for adequate phase-to-amplitude mixing (Talbot effect)



High-Order Wavefront Sensing and Control (HOWFS)

- To correct quasi-static speckles:
  - Estimate and control starlight directly in **focal plane.**
  - Use **science camera as WFS** to estimate all aberrations.
- **Estimation + Control** (= Correction) is **iterative**:
  - Model errors, estimation errors, nonlinearities

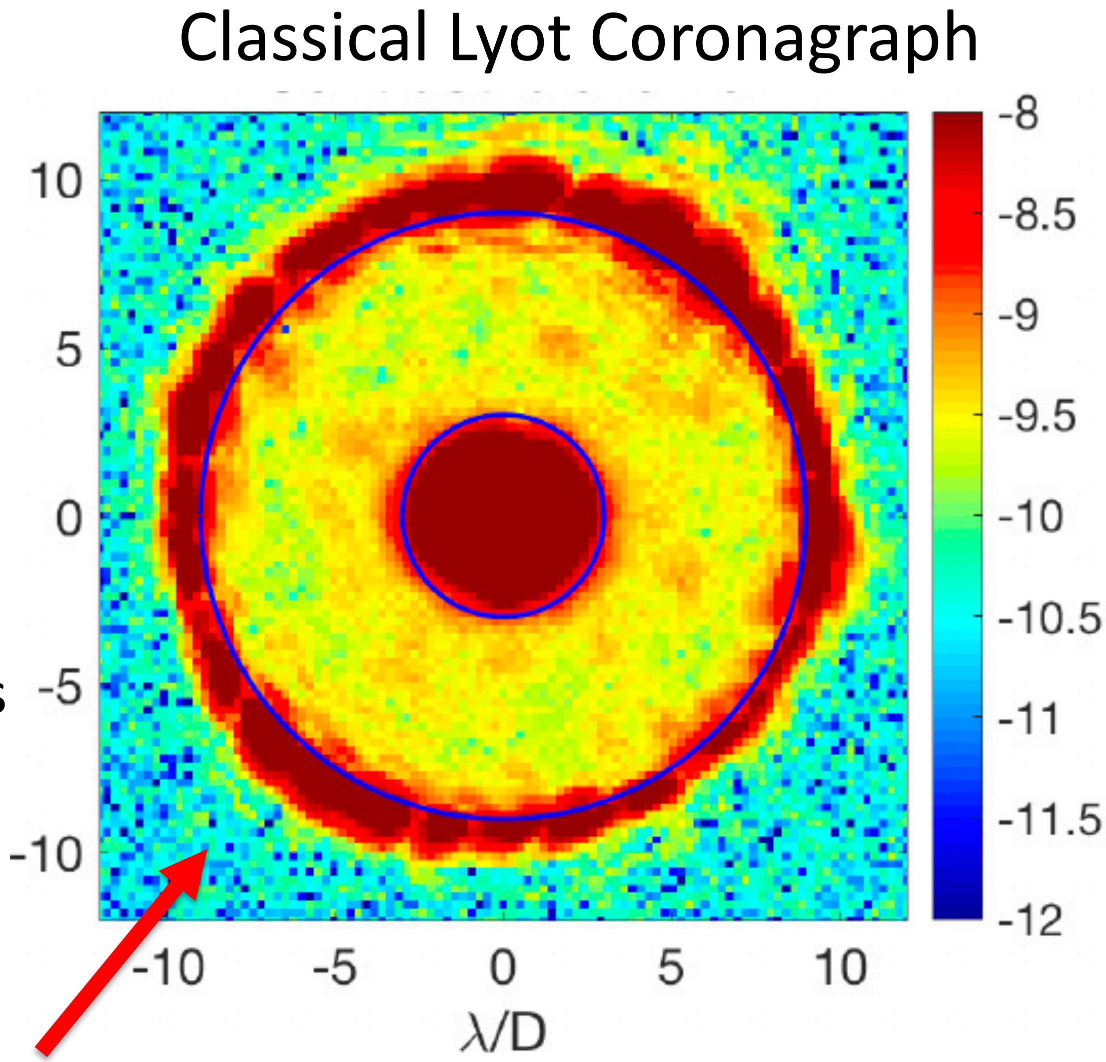
x_k = state
$\hat{x}_k$ = state estimate
z_k = measurement
u_k = control vector



Closed Loop Laboratory Example

10% band centered at 550 nm.

Average Contrast is 3.8e-10.



IWA (3 λ/D) set by coronagraph design, OWA (8 λ/D) set by number of actuators on DM.

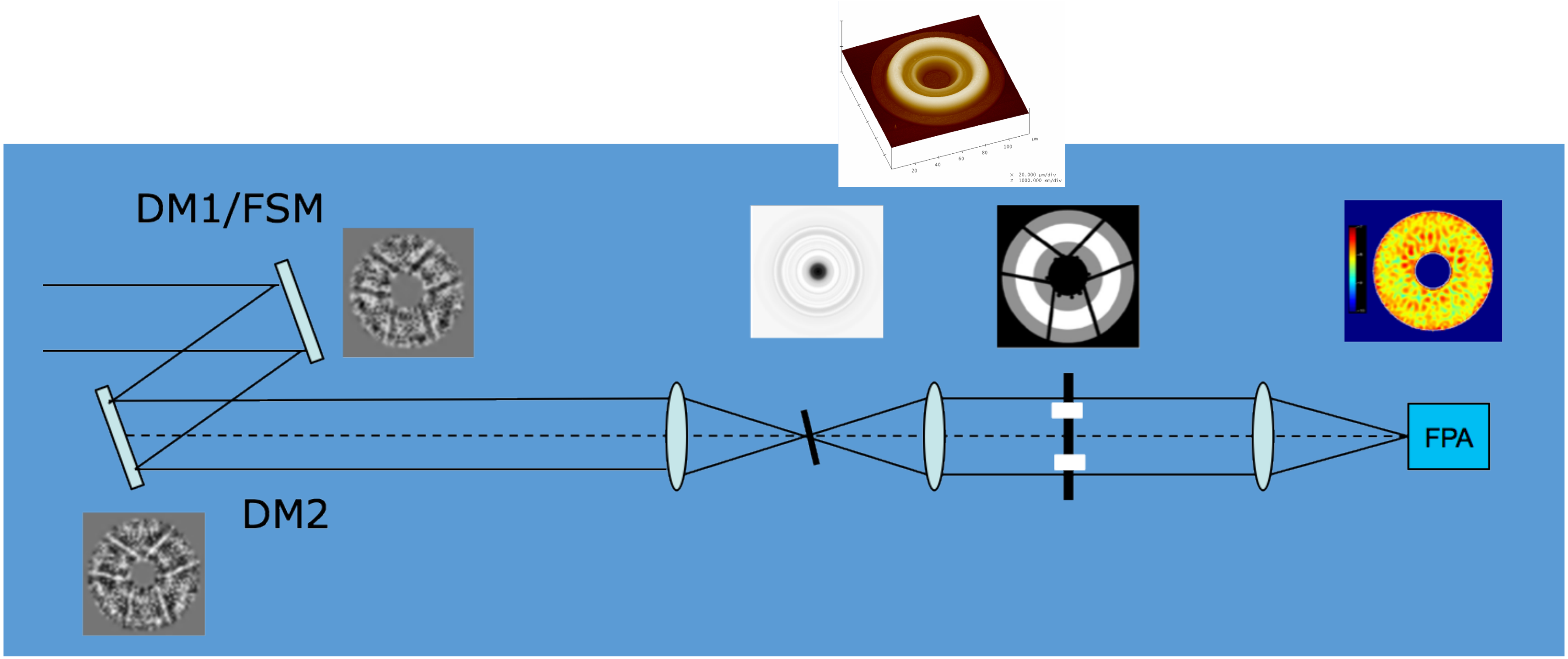
Image is blocked by a focal plane stop outside the OWA.

Seo, B-J. et al. 2019

With DMs, it is possible to use them to generate the coronagraph contrast, improving contrast, throughput, and IWA.

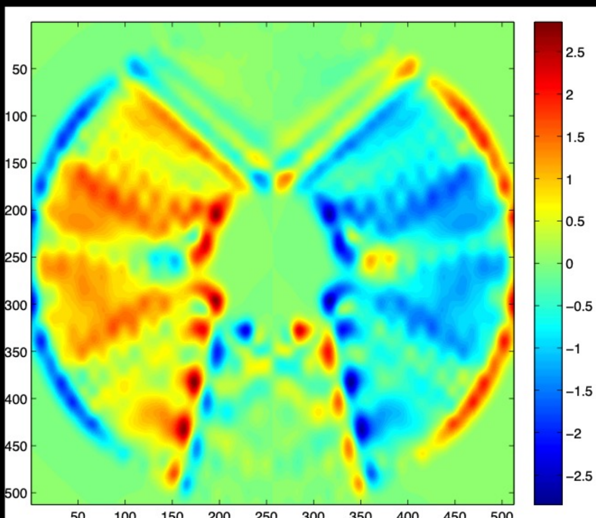
These are called Hybrid coronagraphs.

Hybrid Lyot (Roman Coronagraph)

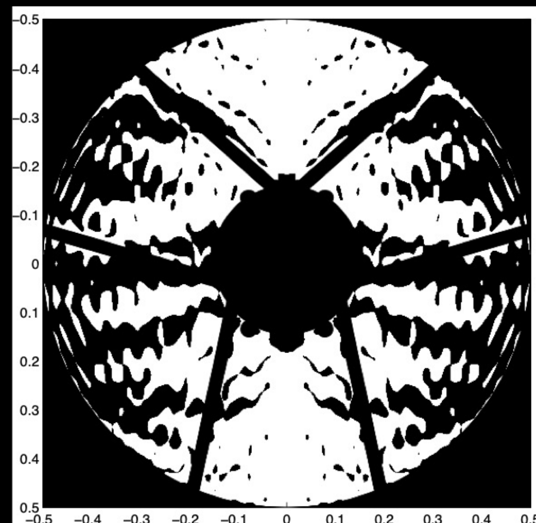


Increases throughput, maps out obstructions, and broadens bandwidth from classical Lyot.

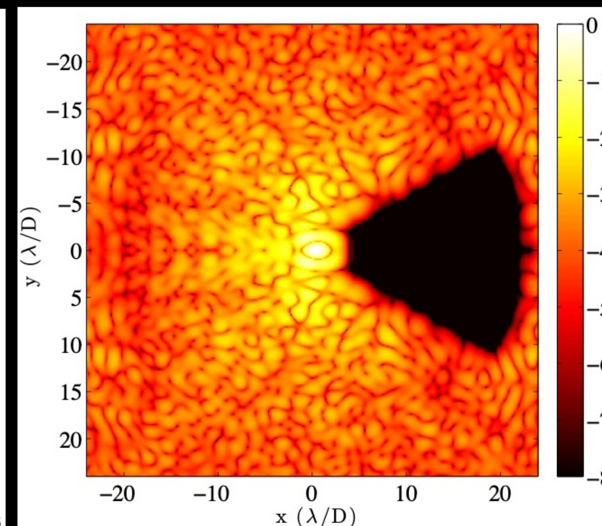
Hybrid Coronagraphs



DM Setting



Shaped Pupil

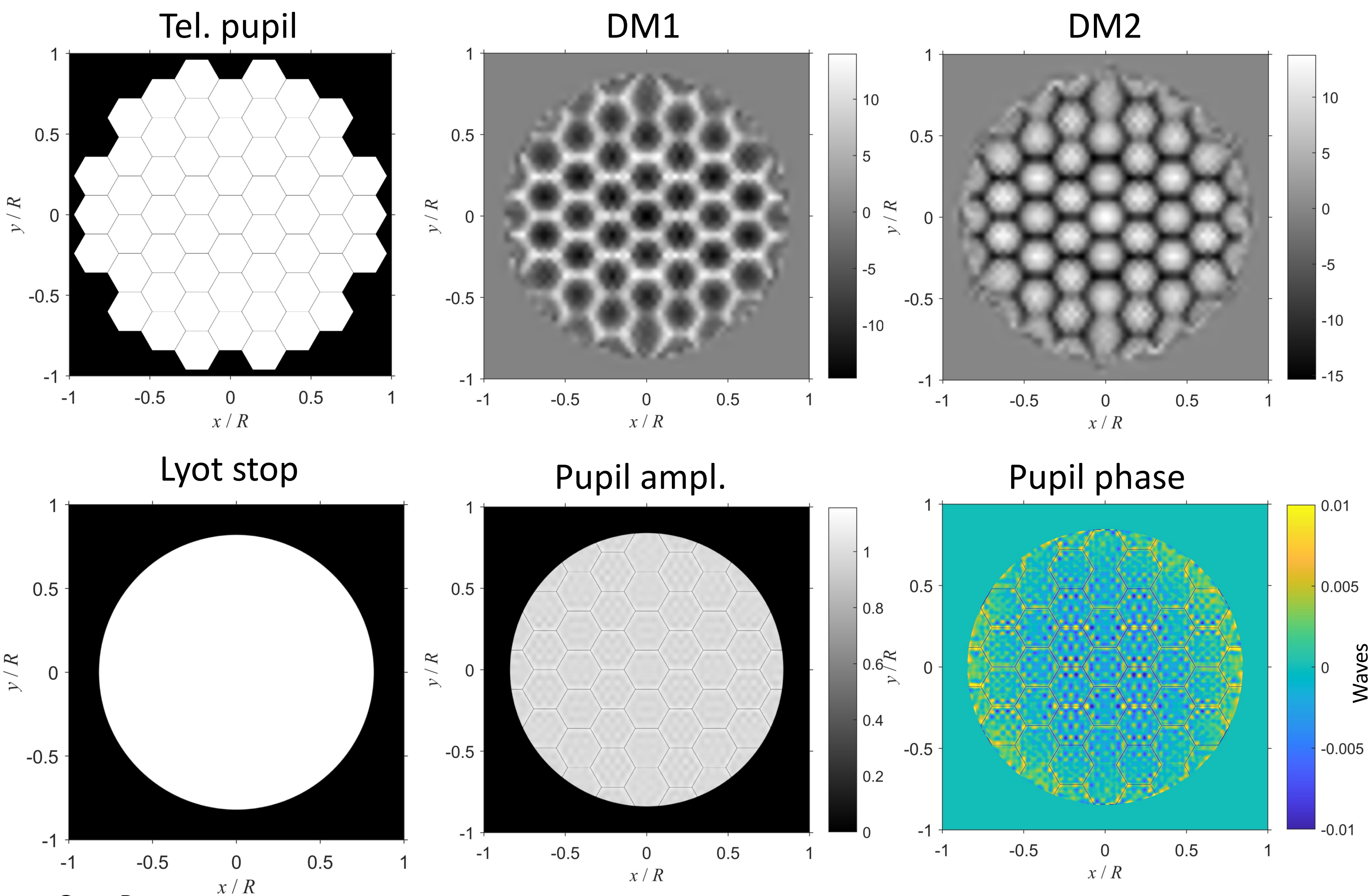


One-Sided Dark Hole

- Contrast: 5x10^{-9}
- Transmission: 61%
- Stroke: 0.91 \lambda
- IWA: 4 \lambda/D
- OWA: 22 \lambda/D

Riggs, et al. (2014)

Segmented Pupil
DM Apodized Vortex Coronagraph



Courtesy Gary Rouane.

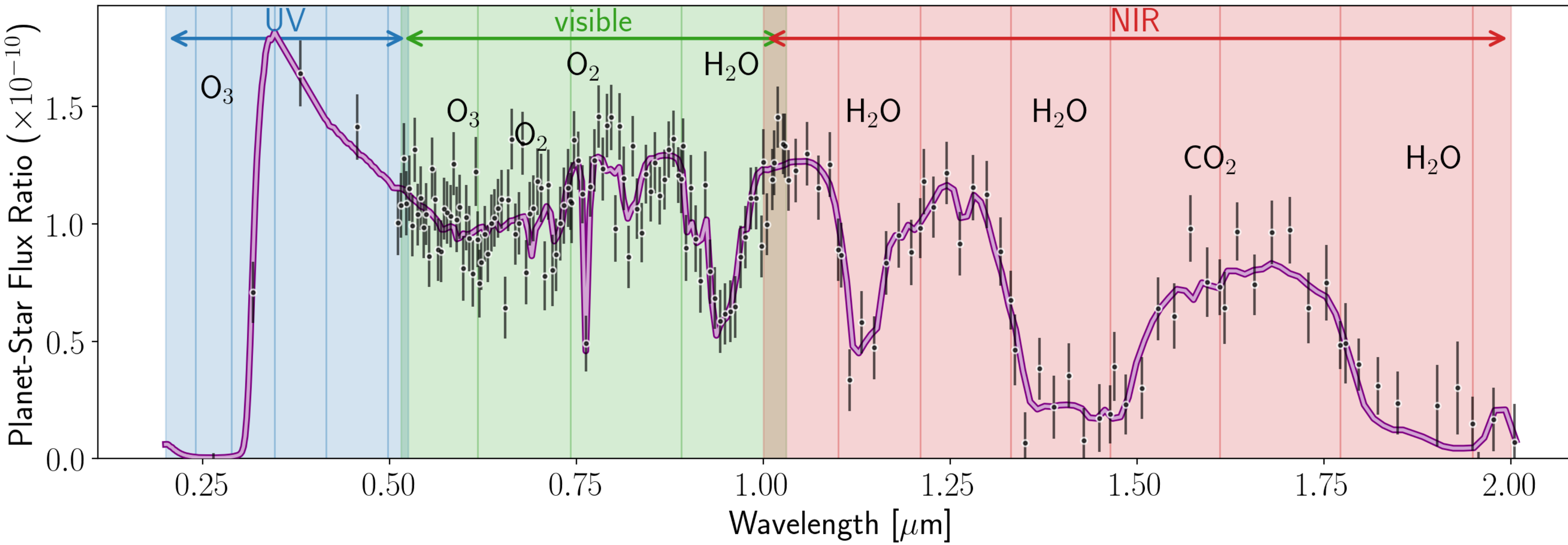
Coronagraph Metrics

- Contrast: The ratio of the peak of the stellar point spread function to the halo at the planet location.
- Inner Working Angle: The smallest angle on the sky at which the needed contrast is achieved and the planet is reduced by no more than 50% relative to other angles.
- Throughput: The ratio of the light in the planet PSF two the nominal telescope PSF after high-contrast is achieved.
- Bandwidth: The wavelengths at which high contrast is achieved.
- Sensitivity: The degree to which contrast is degraded in the presence of aberrations.

Coronagraph performance also differs depending upon aperture (monolith vs. segmented, off-axis vs. on-axis)

Performance – Bandwidth

The LUVOIR Report



Coronagraph must suppress starlight over bands from 10% to 20% for efficient spectroscopy at varying resolution.

Spectral resolution possible is determined by throughput and properties of detector (read noise, dark current, cosmic rays, stability) and type of spectrometer (IFS vs. Pointed).

Performance – Bandwidth

Coronagraphs differ in degradation with bandwidth.

Pupil based coronagraphs generally insensitive to bandwidth but lose iwa.

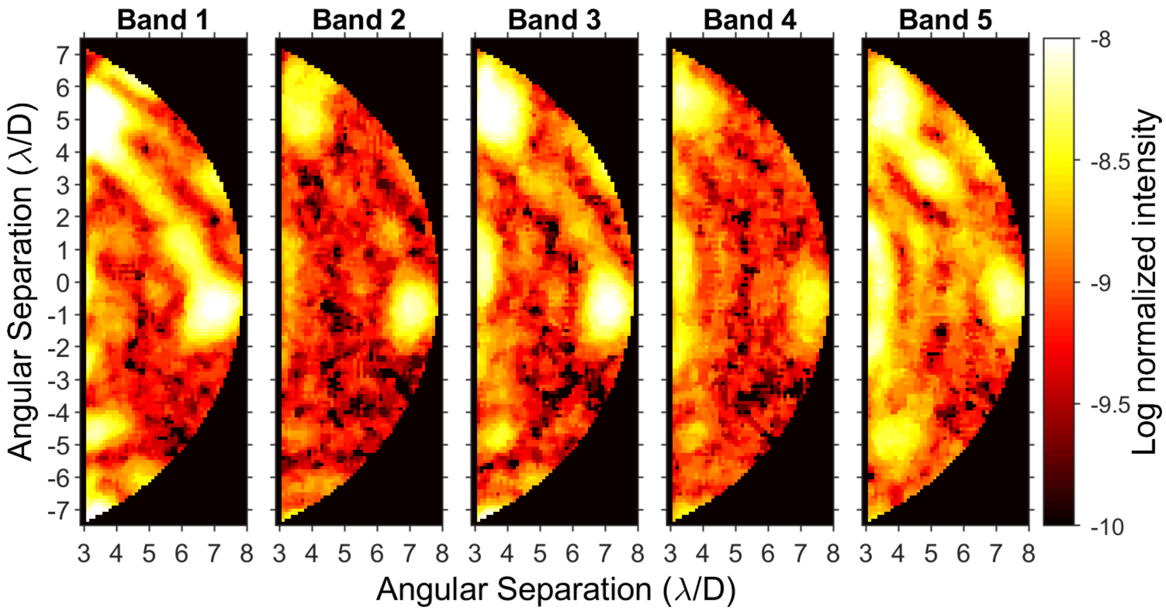
Focal plane coronagraphs generally have limited bandwidth due to spot size, though can be optimized via phase and amplitude variation. (e.g., HLC)

Focal plane phase varying coronagraphs are wavelength independent if broadband spot can be manufactured. (e.g., VVC)

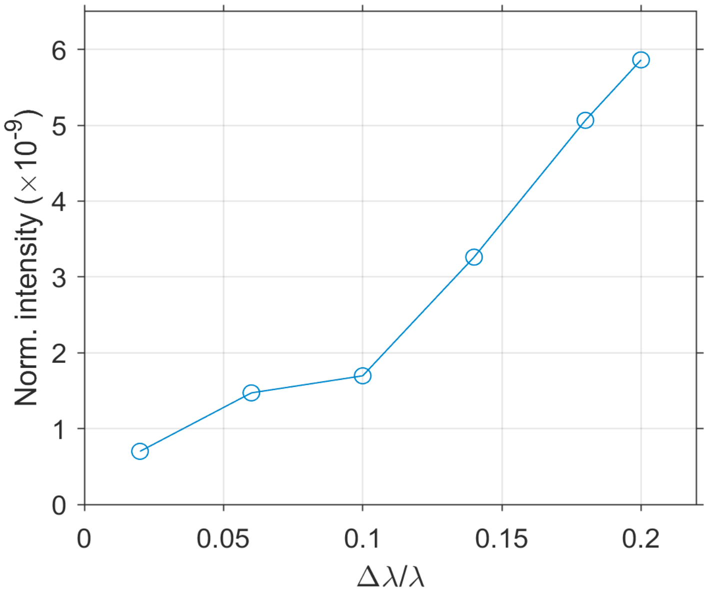
Main limiter of bandwidth is wavefront control via DMs.

Pueyo et al. 2007 show that in a multi-optic system, phase and amplitude errors can be written in a power series in $1/\lambda$. A single DM corrects $1/\lambda$ phase. Two DMs correct for lambda independent amplitude and $1/\lambda$ phase. The remaining terms set the bandwidth of the correction.

Ex.: VVC Lab Result



Raw normalized intensity images obtained in five 2% sub-bands with a VVC4 operating on an unobscured circular aperture. DM optimized for 10% band around 650 nm.



Spatial average of normalized intensity measured over the dark hole with the same VVC4 set-up, but optimizing the DM settings for spectral bandwidths ranging from 2% to 20%.

Ruane, G. et al. 2022

Some wavelength challenges

Recall from Giada Arney's talk that there is a strong desire to get spectra in both the UV (<300 nm) and NIR (>1000 nm) to avoid false positives.

Both are challenging for Coronagraphy!

- Near IR requires a coronagraph with very small IWA in lambda/D to reach habitable zone
- UV has very low throughput due to low reflectivity and large number of optics
- Wavefront control in UV is challenging, requiring high resolution DMs
- Low noise, stable, high QE detectors required for both

This is an incomplete list. Current work is directed at meeting these challenges.

Performance – Sensitivity

Coronagraphs differ in their sensitivity to optical aberrations and stellar diameter

Dynamic aberrations

- Fast low-order variations (e.g., tip/tilt, jitter)
  - Low-order wavefront sensing and control (LOWFS)
- Slow, quasi-static aberrations (actuator drift, thermal creep)
  - Sets picometer level stability requirements on observatory

Static aberrations

1. Low-order aberrations – **global** Zernikes
2. Segment-level aberrations – **segment** Zernikes
   a) Uniform segments
   b) Randomized segments
3. Mid-spatial frequency aberrations - **PSD** errors
4. Lateral beam shear
5. Stellar diameter

Juanola-Parramon et al.

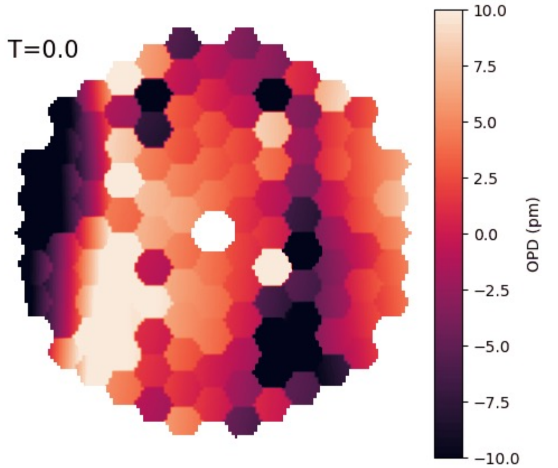
Ex.: Fast Variations

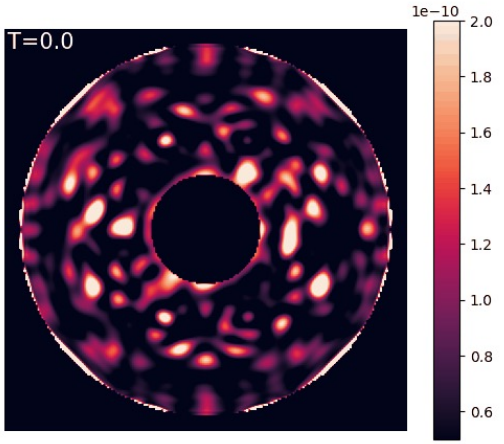
- LMC performed a **Finite Element Model** of the telescope and spacecraft structural dynamics. It takes into account:

  Rigid body motion of the primary mirror segments and subsequent optics relative to each other

  Dynamic interaction of flexible structures

  Disturbances from the multi-stage pointing control system, primarily reaction wheels.

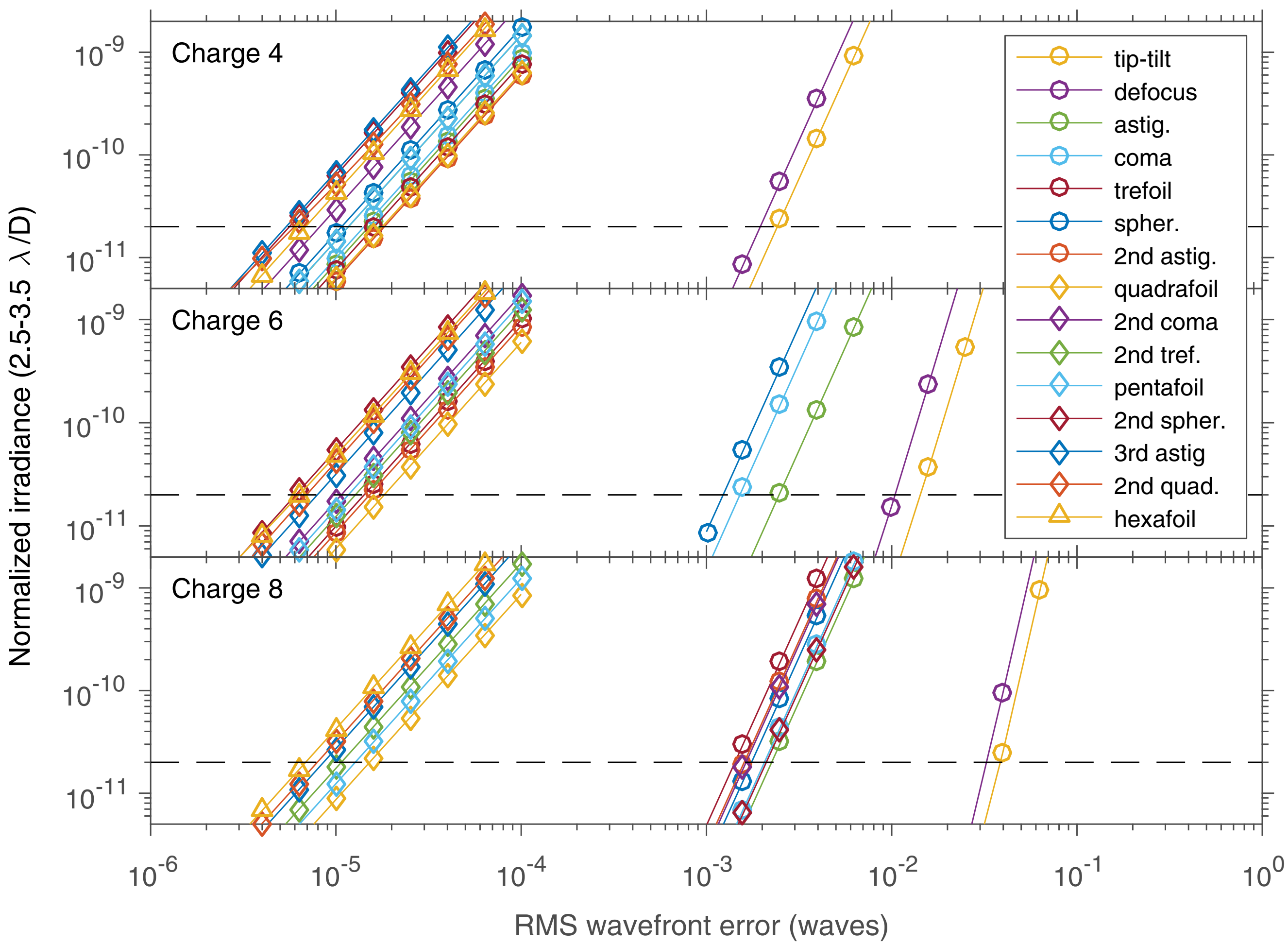




Juanola-Parramon et al.

Performance – Sensitivity

Ex.: Low-Order Zernike Aberrations



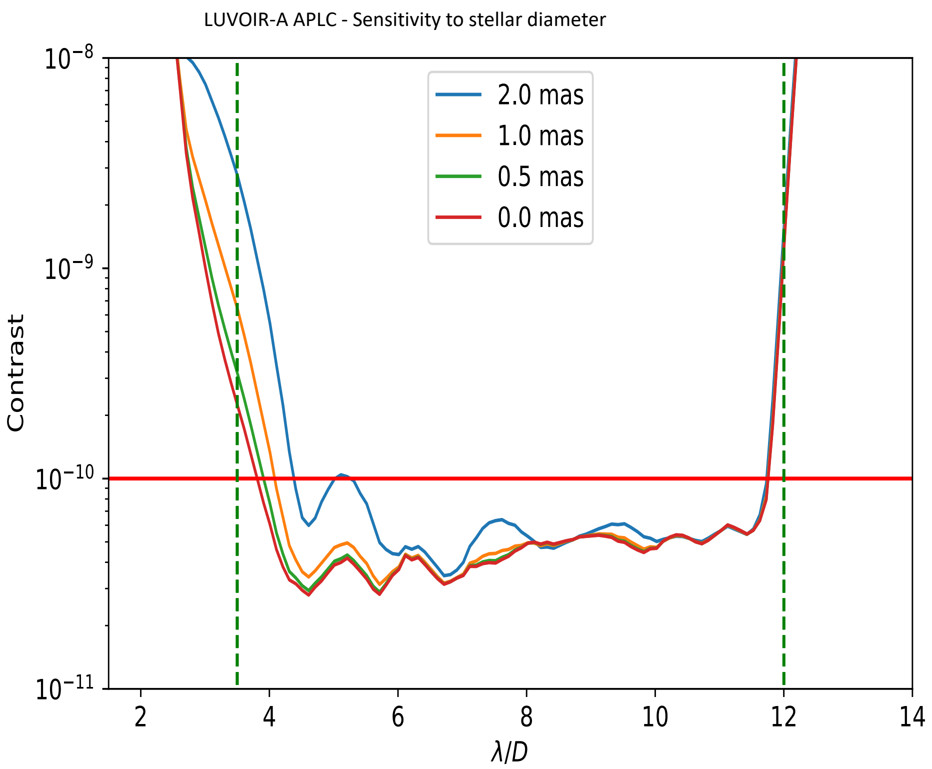
VVC

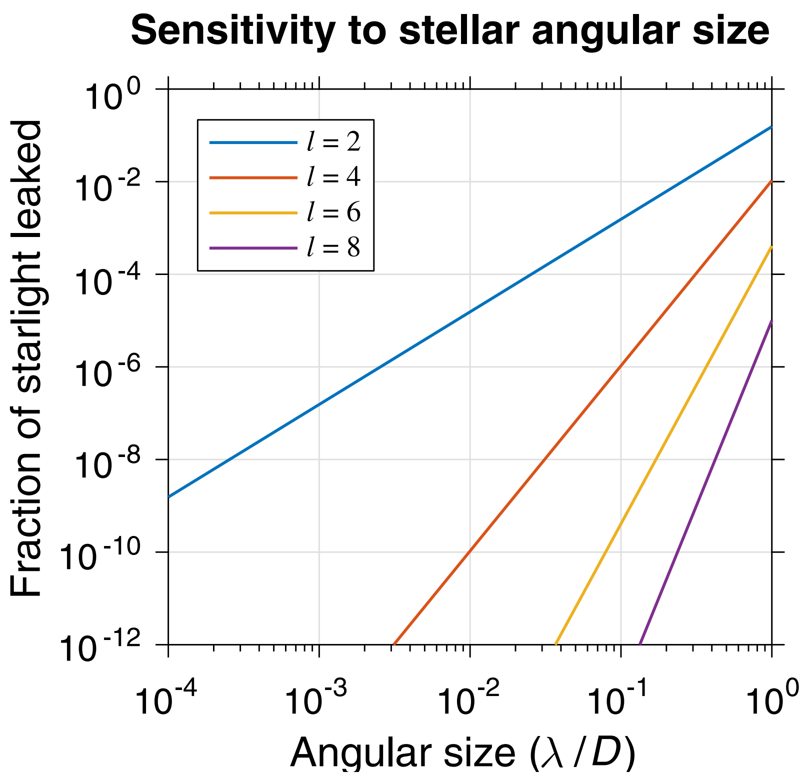
Higher charge less sensitive at cost of throughput and iwa. (See previous slide on throughput.)

Ruane et al.

Performance – Sensitivity

Stellar Diameter – sets limits on close stars





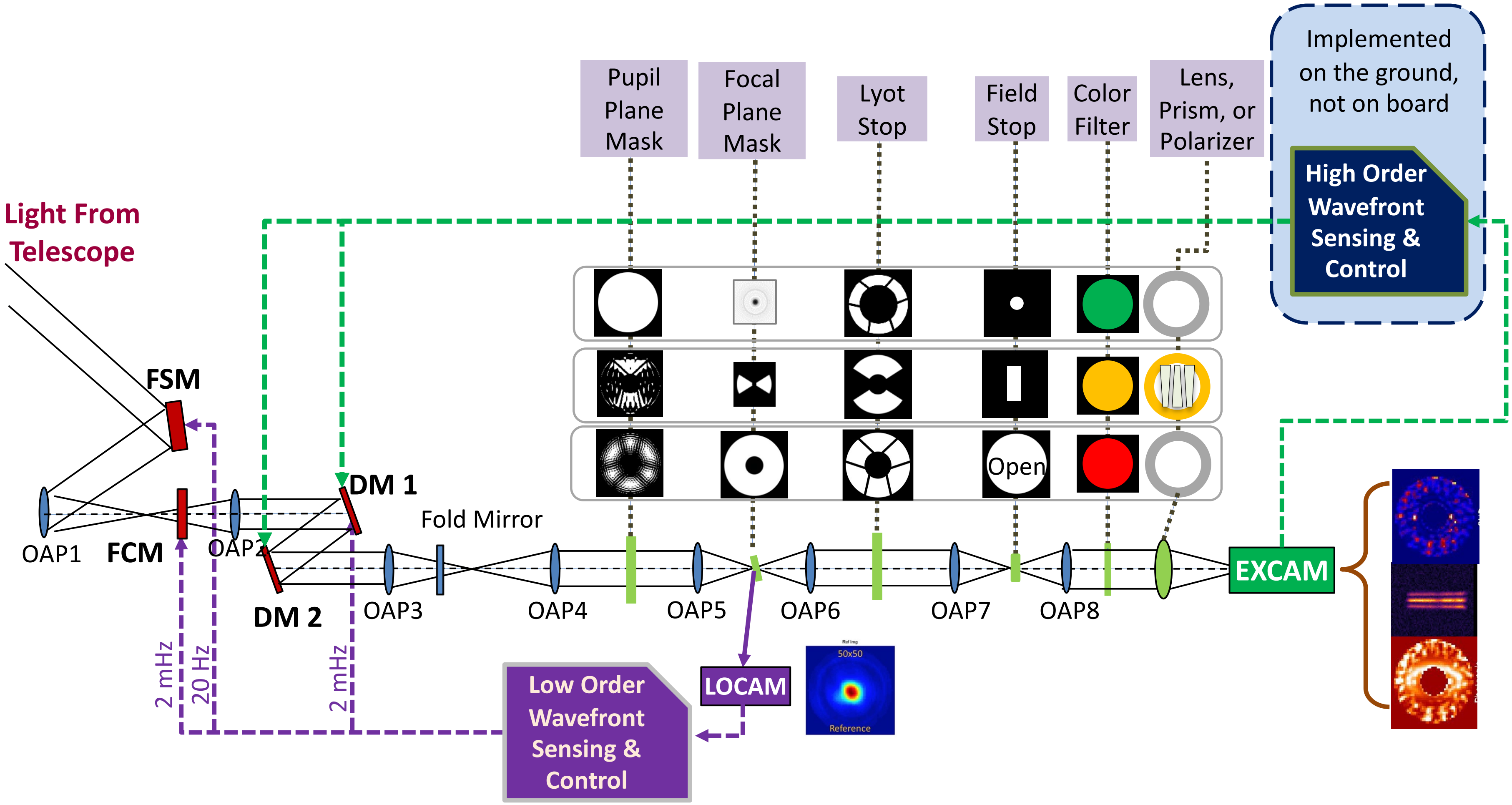
APLC

Juanola-Parramon et al.

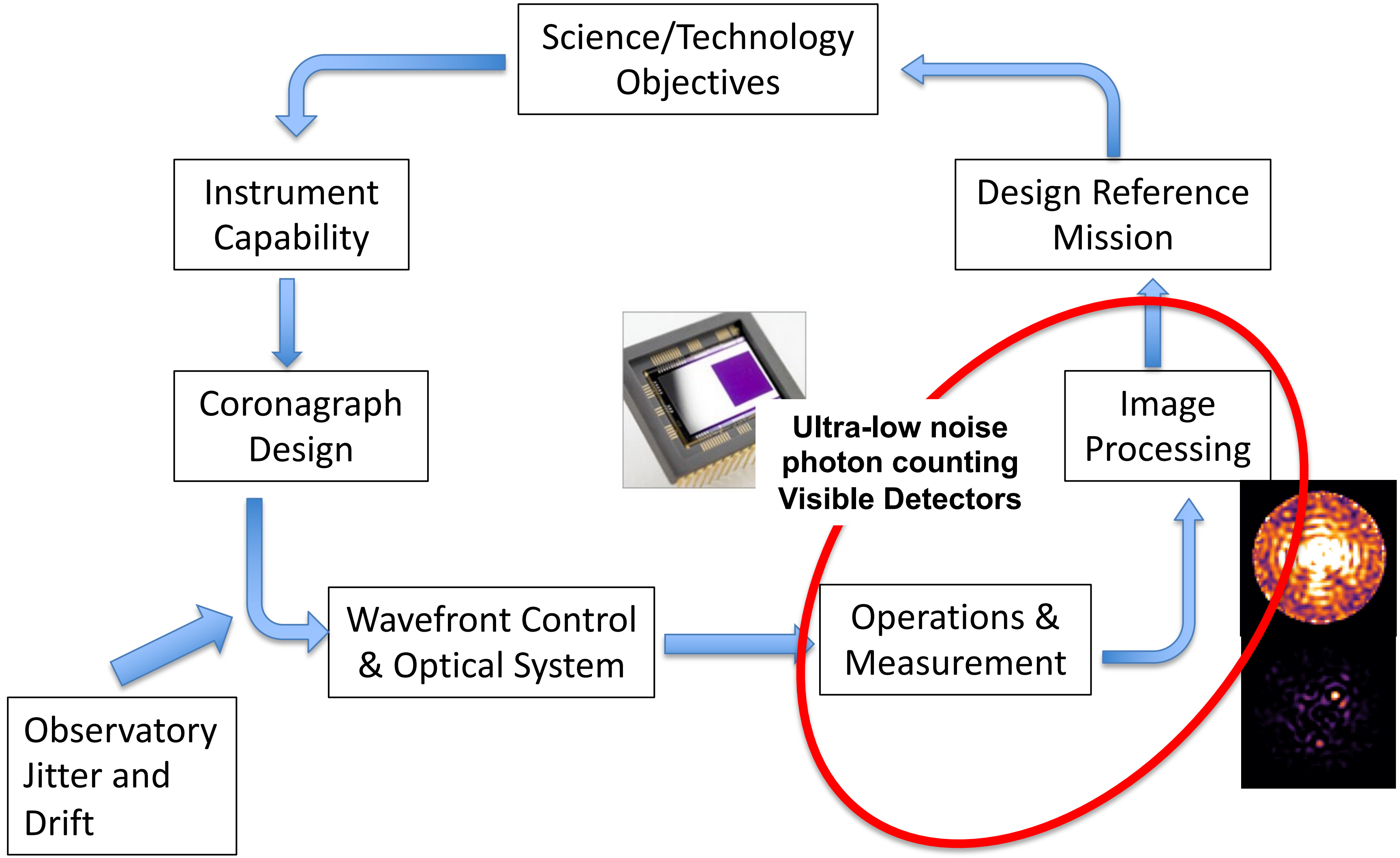
VVC

Ruane et al.

Putting it all together

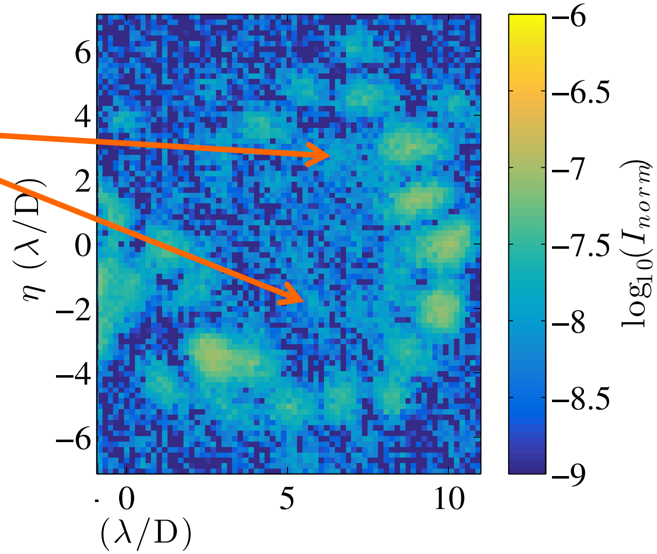


The Roman Coronagraph System


Science/Technology Objectives
Instrument Capability
Coronagraph Design
Wavefront Control & Optical System
Observatory Jitter and Drift
Ultra-low noise photon counting Visible Detectors
Operations & Measurement
Image Processing
Design Reference Mission

How is planet differentiated from residual speckles?

Planets or speckles?



JPL HCIT lab data

Subtract the remaining PSF to remove speckles and reveal planet:

1. Reference Differential Imaging (RDI and KLIP)
2. Angular Differential Imaging (ADI)
3. Spectral Differential Imaging (SDI)
4. Coherent Differential Imaging (CDI)

Reference Differential Imaging

RDI: Remove starlight by subtracting a template PSF

Target Star | Template PSF (no companion) | Circumstellar Disk Revealed

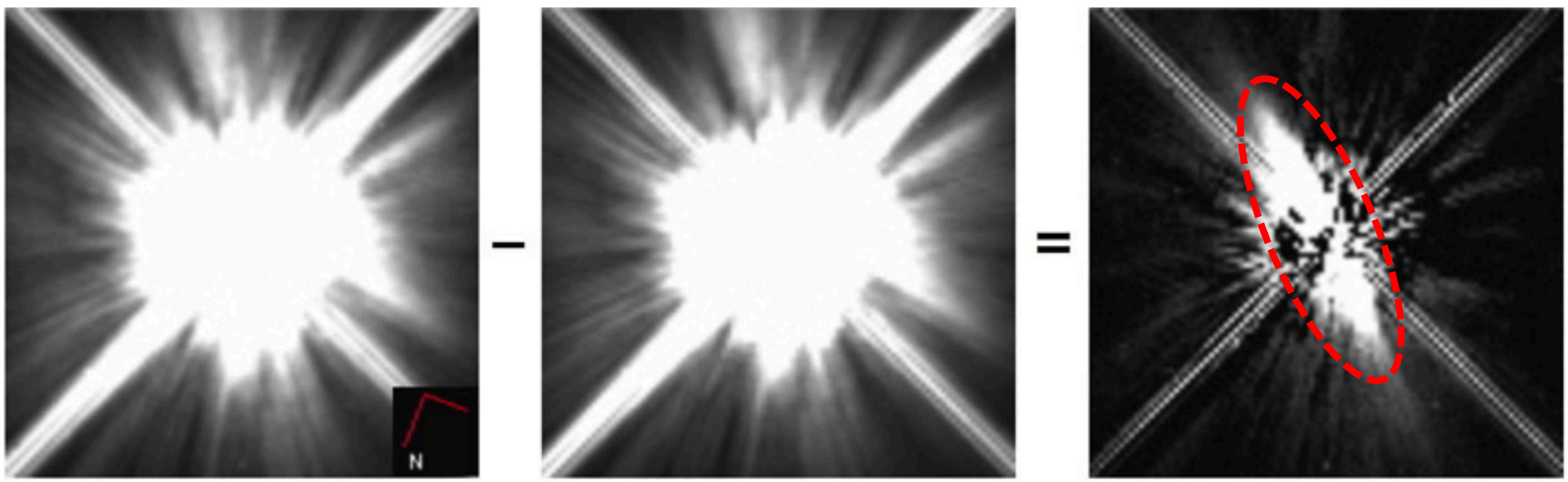


Image Credit: archive.stsci.edu/prepds/laplace/

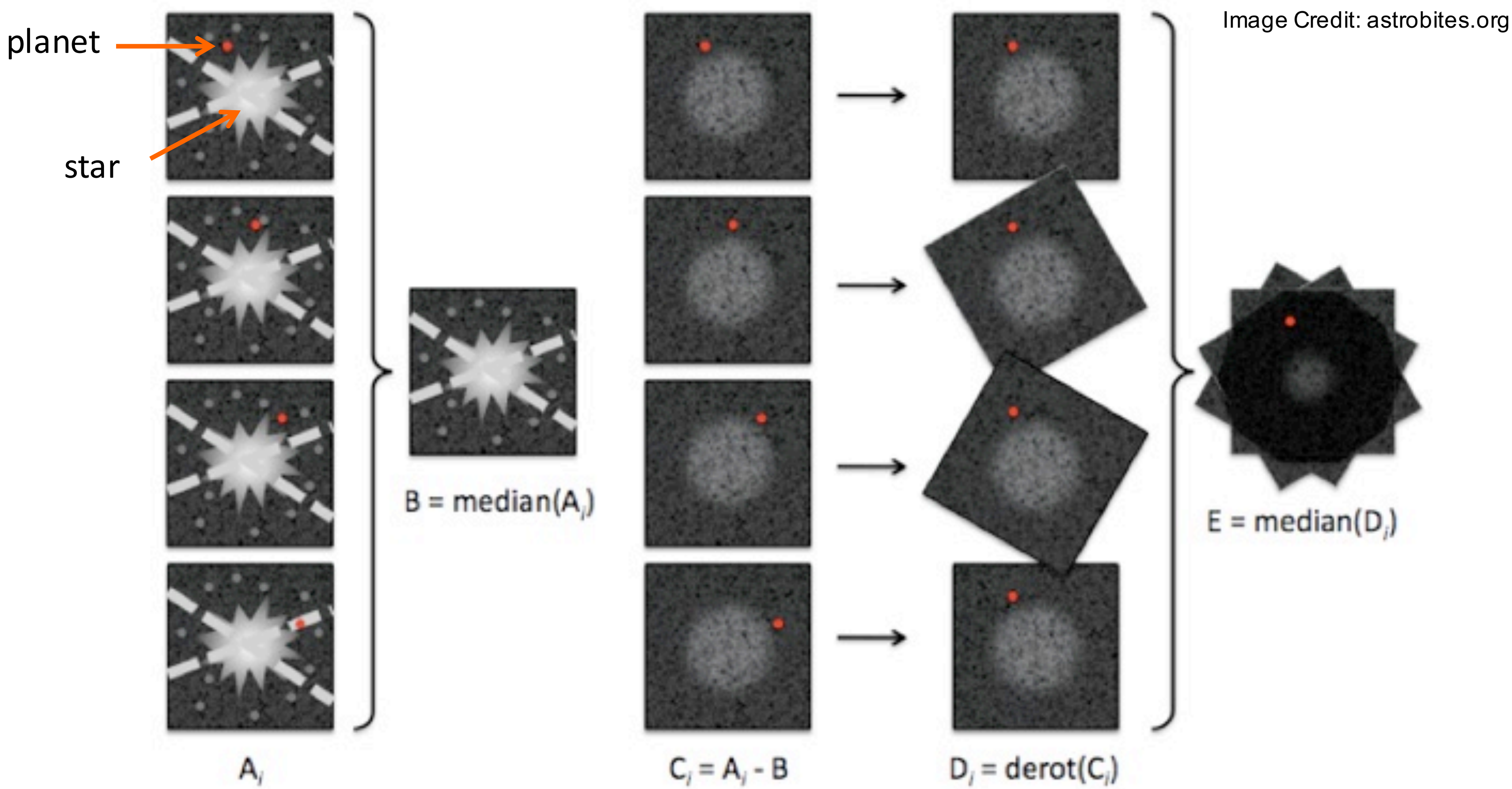
Two variations:

- **PSF Subtraction** (simplest case): Template PSF is directly measured from 1 star
- **Principle Component Analysis** (PCA): Template PSF is a "Franken-image" built from similar parts of many PSFs

Lafrenière+ 2007
Soummer 2012

Angular Differential Imaging

ADI: Take advantage of planet moving w.r.t. stellar speckles during telescope/sky rotation.



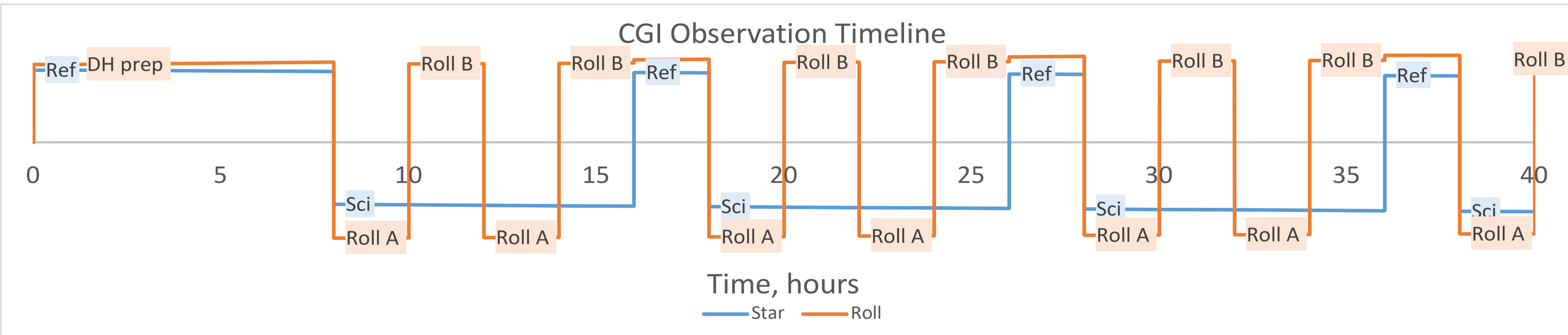
1. Take images at different orientations.

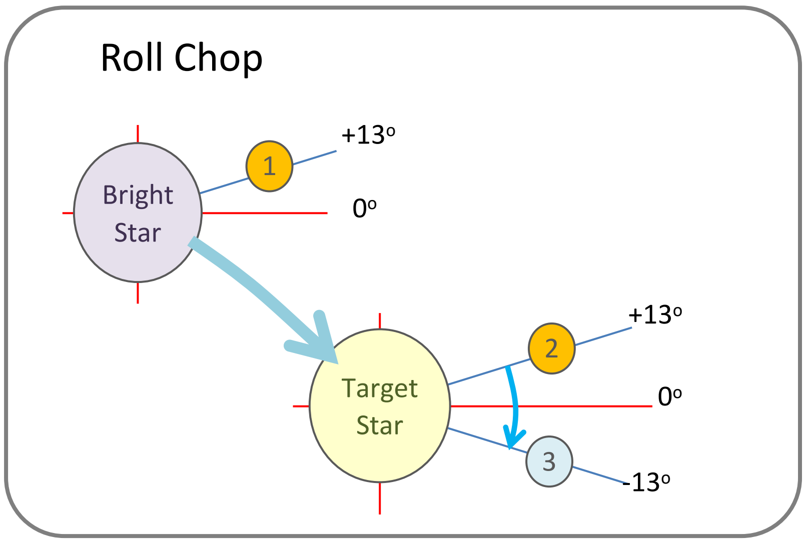
2. Subtract median from all images.

3. Derotate images.

4. Combine images.

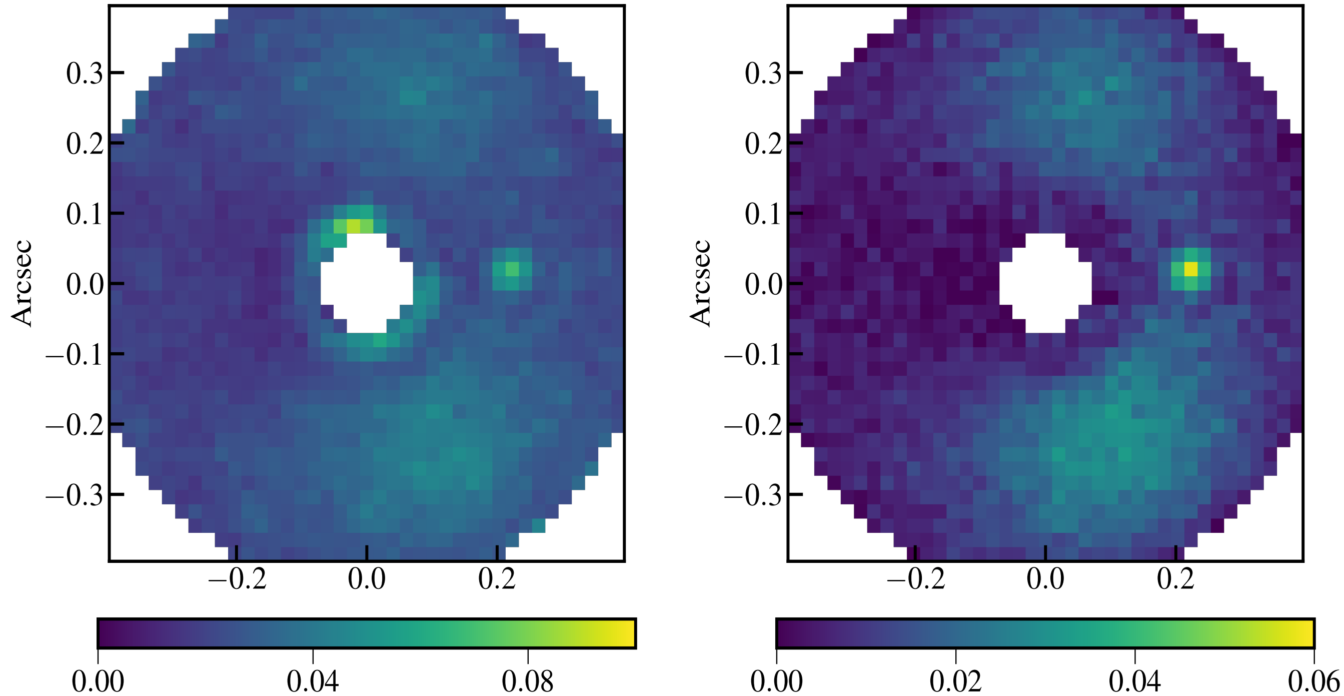
Roman Operational Concept





RDI and ADI enabled through a combination of rolling and chopping to a reference star for a PSF library.

Roman Simulation



Simulated Image of Jupiter Size planet with Exozodi Dust before and after Subtraction.

Courtesy Maxime Rizzo

Integration Time Calculator

$$\mathrm{SNR} = \frac{r_{pl}\, t}{\sqrt{r_n t + \sigma_s^2}}$$

planet rate (r_{pl}), time (t), noise rate (r_n), Speckle variation (σ_s^2)

$$r_{pl} = F_\lambda \Delta\lambda\, \xi_{pl}\, A\, \tau_{pl}\, \eta$$

flux (F_λ), Flux ratio (ξ_{pl}), area (A), throughput (τ_{pl}), QE (η)

Courtesy Bijan Nemati

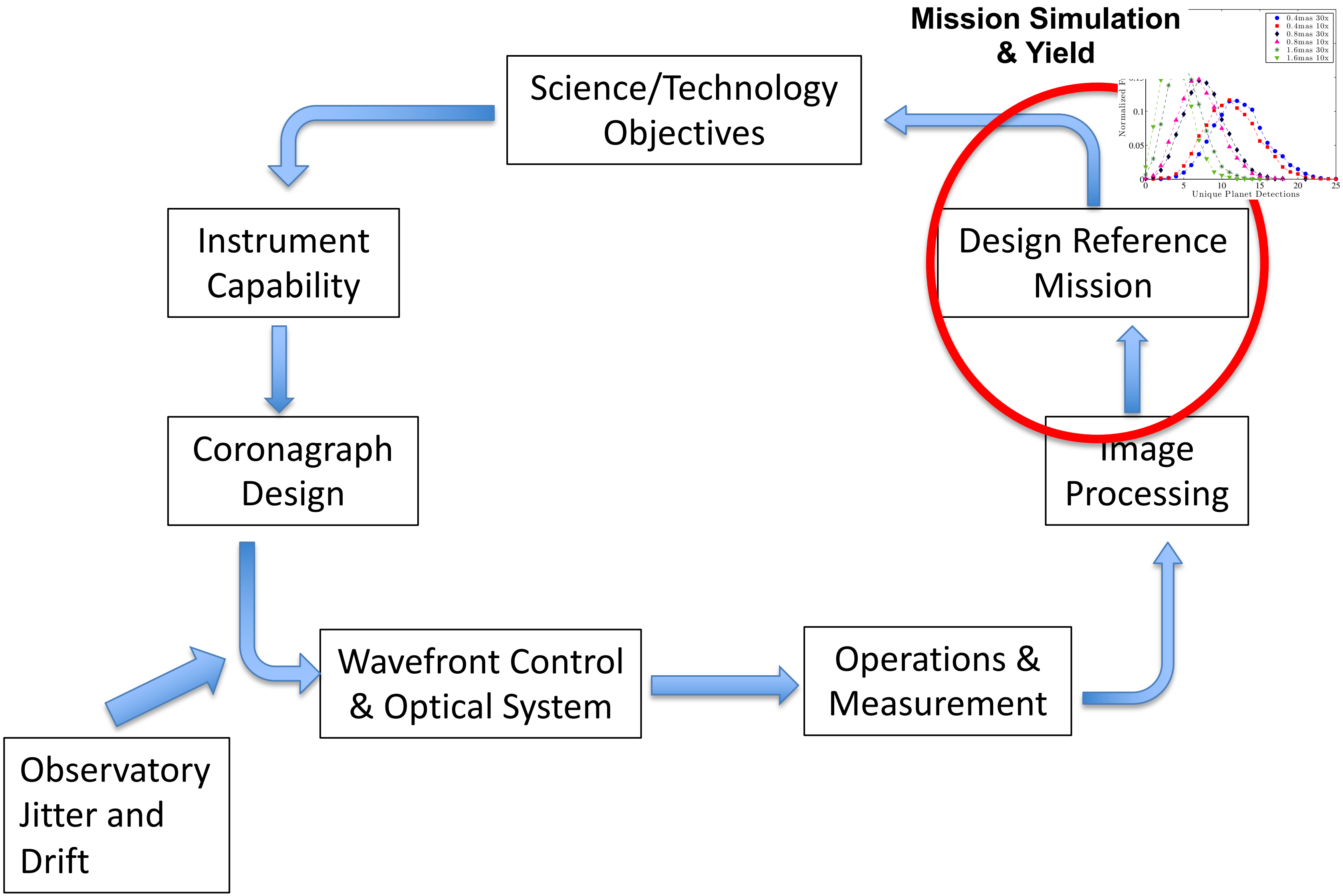
Models for camera noise, shot noise, and speckle stability used to calculate “detection threshold” and integration times.

What is the minimum flux ratio planet we can see with a specified SNR in the allocated integration time?

Likely to be speckle stability limited, sets requirement of contrast stability at better than 10^{-11}.

Integration time calculator used in mission planning tools to determine total mission yield.

Mission Level Analysis



DRM Development

With integration time calculations can build a “Design Reference Mission”, the order and length of observing to determine total science yield.

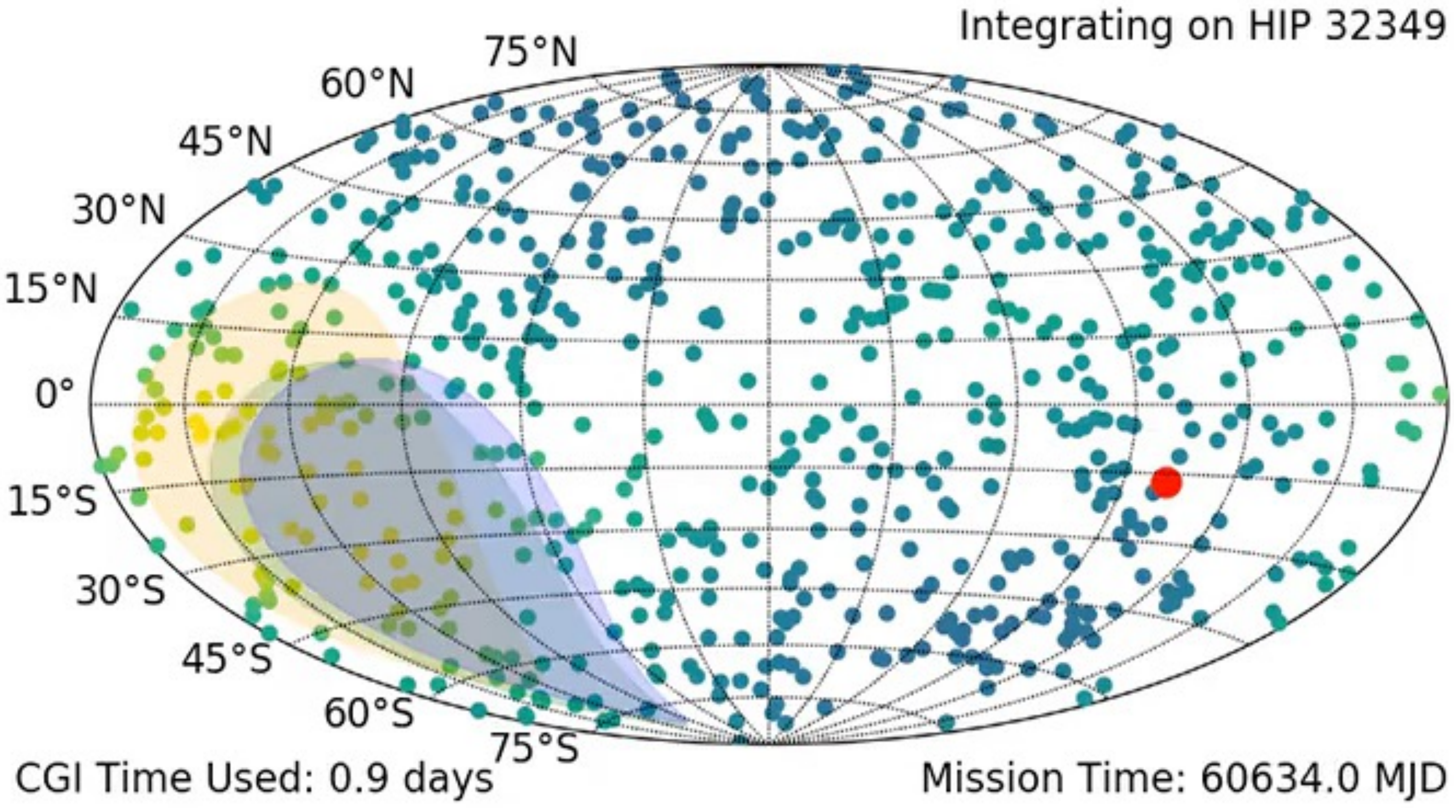
Two general approaches:

- Semi-analytical optimization (Stark)

- Monte-Carlo Mission Builder (ExoSim – Savransky)

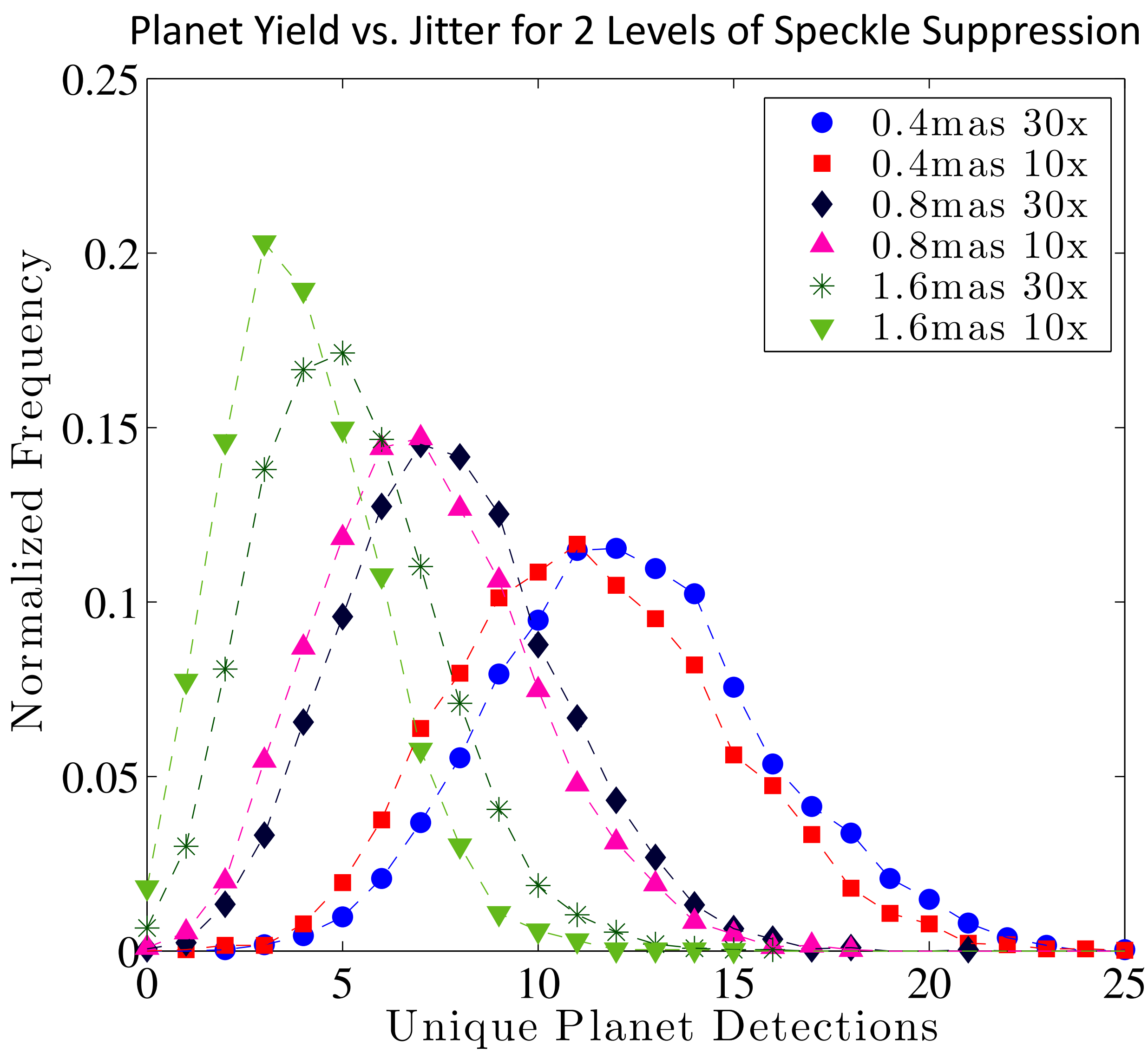
ExoSIM

Courtesy Dmitry Savransky



Courtesy Dmitry Savransky

Sample Output



All of these pieces then come together to produce estimates of capability and potential science yield.

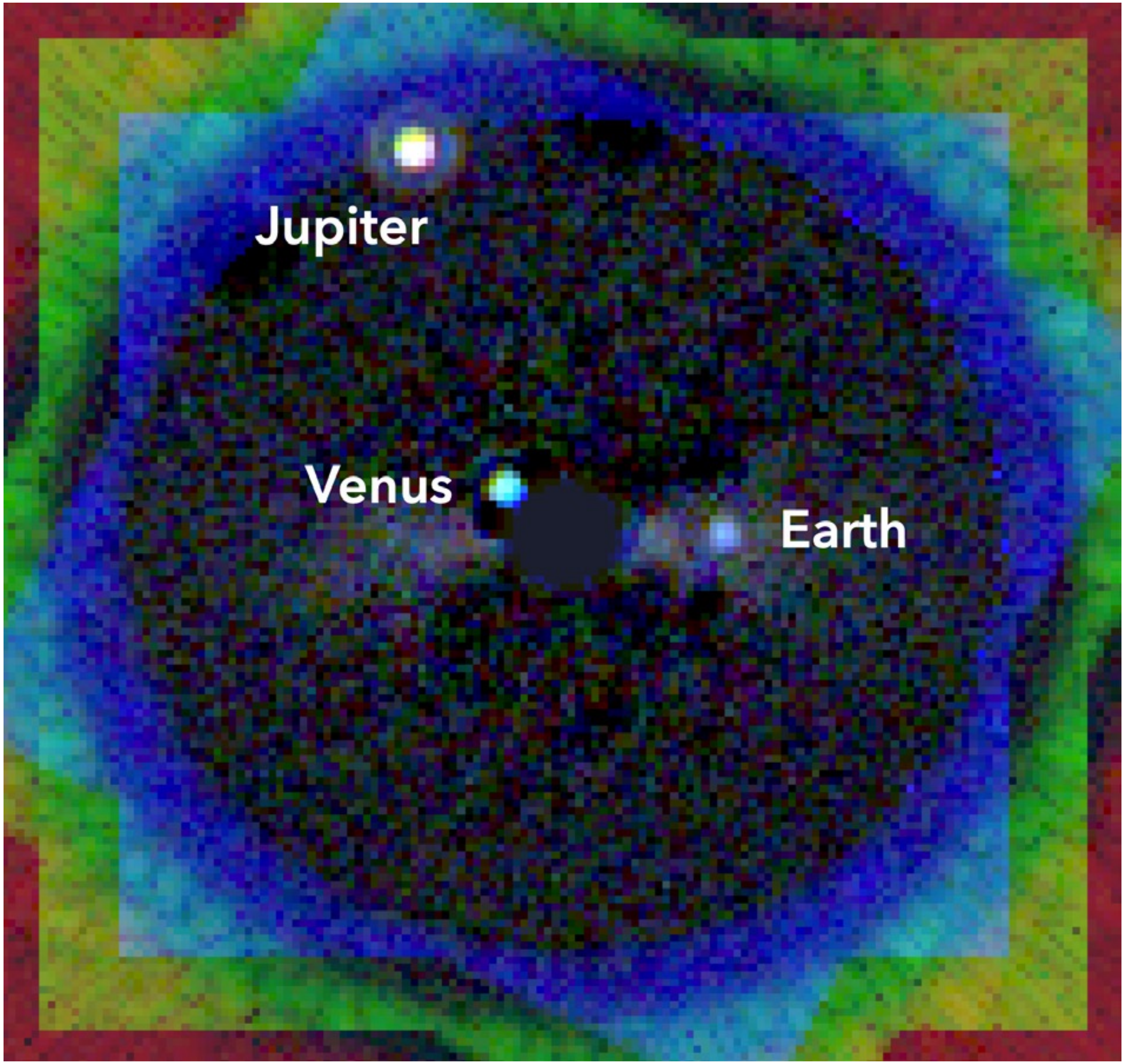
What haven't I addressed?

- Hardware (DMs, Low-noise Detectors, Coronagraph masks, Spectrometer type)
  - Coronagraph Technology Roadmap
- Laboratory results and advances
  - See Mennesson et al.
- Polarization and Polarization Aberrations
  - See Krist et al.
- Recent advanced coronagraph designs (PAPLC, PIAACMC, PIAA vortex, Photonics)
- Advances in WFSC (LDFC, IEFC)

Some challenges ahead

- Achieving 10^{-10} contrast in the lab
- Achieving observatory stability requirements
- Low noise, radiation tolerant, and ultra-stable detectors with high QE
- Increasing throughput (to reach adequate number of systems)
- Throughput limited Spectroscopy
- Advanced coronagraphs reaching into NIR with small iwa

Thank You



LUVOIR Report

Backup Slides

UV Science Case: Additional Notes

➤ Comment on the Impact of Low UV Throughput

- A modest rough estimate based on state-of-the-art puts system throughput (from telescope primary to detector output) in the UV at ~ 3%
- Nature presents us fewer bright stars in the UV compared to vis
- How many viable UV targets are there?
- Preliminary UV target list [E. Mamjek, K. Stapelfeldt, D. Savransky] indicate only a modest loss (26%) of viable targets compared to vis, mostly K stars

# of accessible targets (at GALEX NUV band for O3 detection)

NUV throughput	3%	9%	18%
mag limit (NUV,AB)	35.64	36.84	37.64
N(total stars)	121	150	158*
N(F-type)	66 (max)	66 (max)	66 (max)
N(G-type)	45	55 (max)	55 (max)
N(K-type)	10	29	37
N(M-type)	0	0	0

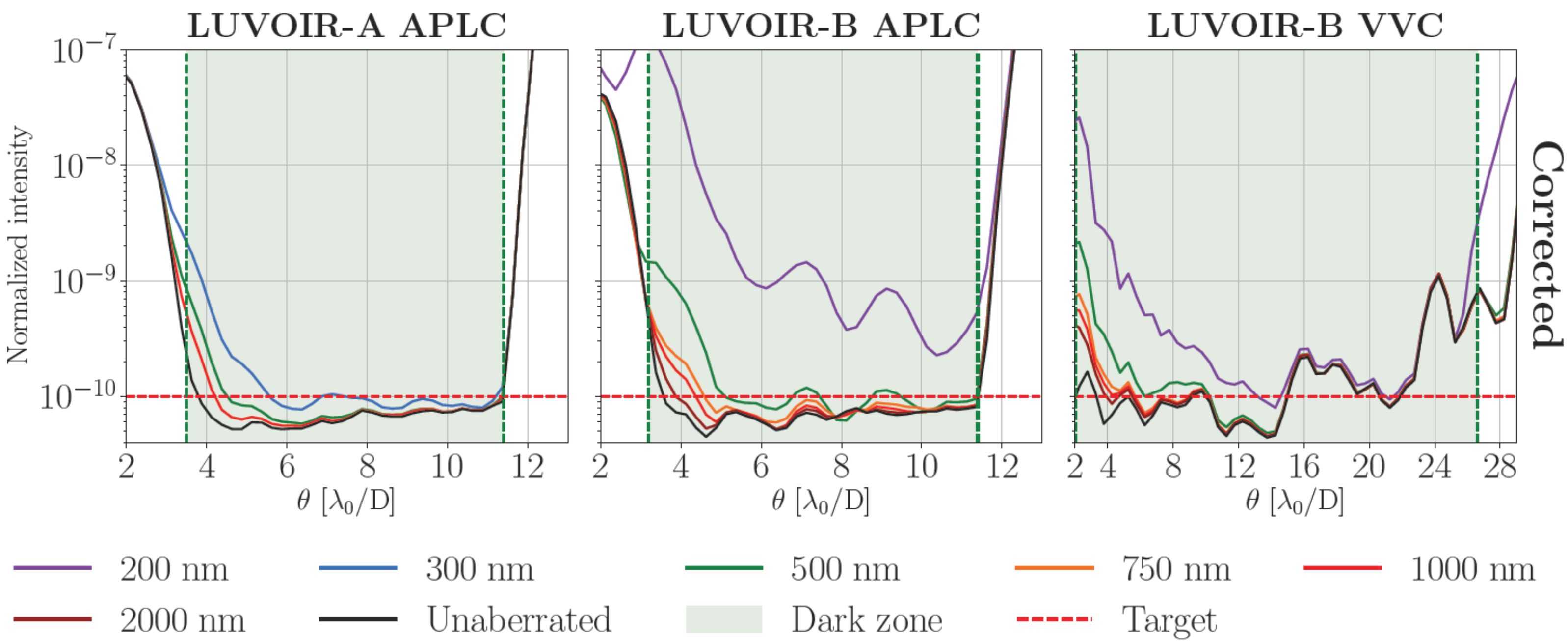
* at 18% throughput, only missing 6 target stars (all K5V-M2V)

Updated *preliminary* estimates from K. Stapelfeldt & E. Mamajek (6/7/23) -
Thanks to D. Savransky for discussions and analysis related to stellar NUV photometry

UV Science Case: Additional Notes

➢ **<u>Comment on Polarization Aberration</u>**

- Contrast degradation due to polarization aberrations is especially pronounced at short wavelengths. However, the effect projects mainly into astigmatism. Thus, degradation occurs at small working angles.
- Mitigation
  - Most UV targets are at large working angles in terms of λ/D. Coronagraphs can be designed to mitigate effects of polarization aberration
  - 100 mas ➔ 12 λ/D @ λ = 250 nm, D = 6 m



Will & Fienup 2019

Courtesy Pin Chen