Introduction to high contrast imaging and coronagraphy
Typical high contrast scene

Circumstellar disk (10^3:1)

Giant planets
(10^{3-9}:1)

Rocky planet
(10^{9-11}:1)

Typical FoV ~1"

[Diagram showing a typical high contrast scene with a circumstellar disk, giant planets, and a rocky planet, along with their respective contrast ratios and a typical field of view (FoV) of approximately 1 arcminute.]
Typical high contrast scene (cont’d)

- Circumstellar disk ($10^{3-5}:1$)
- Giant planets ($10^{3-9}:1$)
- Rocky planet ($10^{9-11}:1$)
- Central star
- Typical FoV ~1″
The need for mild contrast

Thermal emission (young giant planets)

\[ L_{pl}^{eff} = 4\pi R_p^2 \times \sigma T_{eff}^4 \]

From evolutionary models, cooling tracks, function of age

Skemer et al. 2014

- 1600 K (Beta Pic b)
- 1000 K (HR 8799 cde)
- 700 K (max T for cold-start)
- 400 K (max T at 5 Gyr)
- 130 K (Jupiter)
The need for very high contrast

**Reflected light**

\[ L_{pl}^{eff}(t) = \frac{A_{pl}}{4} \times \left( \frac{R_{pl}}{a} \right)^2 \times L_* \times \phi(t) \]

**Thermal emission (old planets)**

\[ T_{eq} = T_* \times \sqrt{\frac{R_*}{2a} (1 - A_{pl})^{1/4}} \]

\[ L_{pl}^{eq} = 4\pi R_{pl}^2 \times \sigma T_{eq}^4 \]
DIRECT IMAGING OF EXOPLANETS

• Direct imaging: taking actual pictures of exoplanets
• Demographics at young ages and large separations
• Direct detection enables detailed characterization:
  • Orbital evolution, dynamical interactions
  • Remote sensing of their atmospheres
  • Formation and disk interaction
DIRECT IMAGING OF EXOPLANETS

• Direct imaging: taking actual pictures of exoplanets

• Demographics at young ages and large separations

• Direct detection enables detailed characterization:
  • Orbital evolution, dynamical interactions
  • Remote sensing of their atmospheres
  • Formation and disk interaction
β Pictoris system (Lagrange et al. 2009)
Fig. 9.—Gallery of imaged planets at small separations (<100 AU). HR 8799 harbors four massive planets (5–10 M\textsubscript{Jup}) on \textit{orbital} distances of 15–70 AU (Marois et al. 2008; Marois et al. 2010b), \(\beta\) Pic hosts a nearly edge-on debris disk and a \(\approx 13\) M\textsubscript{Jup} planet at 9 AU (Lagrange et al. 2009a; Lagrange et al. 2010), a \(\approx 5\) M\textsubscript{Jup} planet orbits HD 95086 at 27 AU (Rameau et al. 2013c; Rameau et al. 2013b), and 51 Eri hosts a \(\sim 2\) M\textsubscript{Jup} planet at 13 AU (Macintosh et al. 2015). Images are from Maire et al. (2015a), Nielsen et al. (2014), Galicher et al. (2014), and Macintosh et al. (2015).
DISKS AS SIGNPOSTS OF PLANETS

Schneider et al. 2014 (HST-STIS)
DIRECT IMAGING OF EXOPLANETS

- Direct imaging: taking actual pictures of exoplanets
- Demographics at young ages and large separations
- Direct detection enables detailed characterization:
  - Orbital evolution, dynamical interactions
  - Remote sensing of their atmospheres
  - Formation and disk interaction
DIRECT IMAGING OF EXOPLANETS

• Direct imaging: taking actual pictures of exoplanets
• Demographics at young ages and large separations
• Direct detection enables detailed characterization:
  • Remote sensing of their atmospheres
  • Orbital evolution, dynamical interactions
  • Formation and disk interaction
MEASURE PLANET SPIN

Snellen et al. 2014
DIRECT IMAGING OF EXOPLANETS

• Direct imaging: taking actual pictures of exoplanets
• Demographics at young ages and large separations
• Direct detection enables detailed characterization:
  • Remote sensing of their atmospheres
  • Orbital evolution, dynamical interactions
  • Formation and disk interaction
BETA PICTORIS B ORBITAL MOTION MOVIE

M. MILLAR-BLANCHAER et al. (2015)
Credit: M. Millar-Blanchaer (Dunlap Institute) & F. Marchis (SETI Institute)
DIRECT IMAGING OF EXOPLANETS

• Direct imaging: taking actual pictures of exoplanets
• Demographics at young ages and large separations
• Direct detection enables detailed characterization:
  • Remote sensing of their atmospheres
  • Orbital evolution, dynamical interactions
  • Formation and disk interaction
INTERACTION WITH HOST DISK

β Pictoris b
Mawet, Absil, Milli et al. 2013
4 PILLARS OF HIGH CONTRAST IMAGING

- Adaptive optics
- Coronagraphy
- Differential imaging
- Post-processing
- Know your star (age, L, distance, proper motion, etc.)!
High contrast imaging in space enabled by stability

HST only had classical Lyot unoptimized coronagraphs, but its unmatched stability benefited, and still benefits all of its imagers: WFC2/3 (nc), STIS (bars), NICMOS (hole), ACS (proper masks)

...so much so that archival data mining is producing wealth of results
Its secret: 3 fine guidance sensors (FGS)

~1 mas pointing accuracy over > 10 min
70’ FoV, R < 17
Not a level-playing field!
Ground-based telescope a clear disadvantage

High contrast imaging through this requires some well thought out architecture!
Do you see the planet?

seeing halo

Oops, forgot something!
Adaptive optics / Wavefront control

Correction DM → (0.01 Hz) → NCPA → Camera/ Image plane sensing

Control real-time
Closed loop (kHz)

Sensing
WFS
Adaptive optics / Wavefront control

Closed loop (kHz)

Correction DM

Control real-time

Sensing

WFS

NCPA

Camera/image plane sensing

(0.01 Hz)
1st gen adaptive optics
2nd gen = more horsepower = more actuators

\[ \frac{n}{2} \frac{\lambda}{D} \]

Diffraction!

Contrast

\[ 10^{-6} \]

Angular separation (\(\lambda/D\))

0 5 \(\lambda/D\)
4 PILLARS OF HIGH CONTRAST IMAGING

• Adaptive optics

• Coronagraphy

• Differential imaging

• Post-processing

• Know your star (age, L, distance, proper motion, etc.)!
“The rareness of total eclipses of the Sun, their short duration and the distances one has to travel to observe them have, for more than half a century, led astronomers and physicists to seek for a method which enables them to study the corona at any time.”

– Bernard Lyot, 1930s
The coronagraph by Bernard Lyot

Then (1939) ...
Coronograph.—The principal defect, that is to say, the light diffracted by the edge of the lens, still had to be overcome. With this object I have made several coronographs, of which the plan is seen in fig. 2 (Plate 11).

The single lens, plano-convex, is shown at A; it forms an image of the Sun on a blackened brass disc at B projecting beyond the Sun by 15" to 20". A field lens C, placed behind the disc, produces an image of the lens A on the diaphragm D in the shadow. The edge of the diaphragm cuts off the light diffracted by the edge of the first lens. A small screen placed in the centre of the diaphragm cuts off the light of the solar image produced by reflection on the surfaces of the lens A. Behind the diaphragm and the screen, protected from the diffused light, a carefully corrected objective E forms an achromatic image of the corona on the plate. The whole apparatus is contained in a tube F, open only during the observations, and coated with thick oil in the interior, to collect the particles of dust. The first lens must be wiped frequently, and with particular care.
And now…
Imaging a star: ray optics description

star → FAR → Lens/telescope → Image

Planet
Star
Adding diffraction: Huygens-Fresnel principle

\[
E_0(x, y) = \frac{1}{j\lambda} \int \int_{\Sigma} E_1(\xi, \eta) \frac{\exp(jkr_{01})}{r_{01}} \cos \theta d\xi d\eta
\]

\[
r_{01} = \sqrt{z^2 + (x - \xi)^2 + (y - \eta)^2}
\]

Too complicated => approximations
Fresnel and Fraunhofer approximations

Fresnel (near field)

\[ E(x, y) = \frac{e^{jkz}}{jkz} \int \int_{-\infty}^{\infty} E(\xi, \eta) e^{jk \frac{z}{2z} [(x-\xi)^2 + (y-\eta)^2]} \, d\xi \, d\eta \]

Fresnel #

\[ \frac{R^2}{\lambda z} \sim O(1) \quad x, y \ll z \]

Fraunhofer (far field)

\[ E(x, y) = \frac{e^{jkz} e^{jk \frac{z}{2z} (x^2 + y^2)}}{jkz} \int \int_{-\infty}^{\infty} E(\xi, \eta) e^{-j \frac{2\pi}{\lambda z} (x\xi + y\eta)} \, d\xi \, d\eta \]

\[ \frac{R^2}{\lambda z} \ll 1 \]

= Fourier Transform !

or a lens/telescope
Useful examples

- circ
  - $2J_1(r)/r$
- rect(x)rect(y)
  - $sinc(x)sinc(y)$
- Gaussian
  - Gaussian
- FT
- FT
- FT
Star and planet with diffraction

\[ P_c(\omega) = \left| \mathcal{F}\{ E_i A(x) \} \right|^2 \]

On-Axis Point Spread Function

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

Off-Axis Point Spread Function
What is a coronagraph?

• “A coronagraph is an optical device designed to suppress (or strongly attenuate) the on-axis coherent starlight while allowing the off-axis planet (or circumstellar disk) light to transmit through.”

• **Important definitions:**

  • **Contrast:** The ratio of the planet light to the star light.

  • **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:** fraction of planet light in your photometric aperture.

  • **Bandwidth:** The wavelengths at which high contrast is achieved.

  • **Sensitivity:** The degree to which contrast is degraded in the presence of aberrations.
Lyot coronagraph

Telescope Pupil
Evenly Illuminated

Image is made (top)
And occulted (bottom)

Pupil is reimaged (top)
And partially blocked (bottom)

The Final image after
Coronagraph has only
1.5\% of the original
Starlight.

Occulting Spot

Lyot Stop

Credit: Sivaramakrishnan
Lyot coronagraph cont’d: step by step
Band-limited Coronagraphs (Kuchner & Traub 2002)

Fig. 1.— One-dimensional coronagraph with a Gaussian image mask, examined in the pupil plane. First an incoming field hits the primary aperture, the image mask. The mask ITF multiplies the image intensity; in other words, the conjugate of the image mask ATF \( b \) becomes convolved with the aperture function \( a \). The result, \( c \), can be passed through a Lyot stop, \( d \), leaving the final field \( e \).

Fig. 2.— Ideal band-limited coronagraph with the same aperture \( a \) as the coronagraph in Figure 1. Here the conjugate of the mask ATF \( b \) is identical for spatial frequencies above some cutoff, \( \epsilon D/2\lambda \), so the second pupil field \( c \) for an on-axis source is zero except within \( \epsilon D/2\lambda \) of the aperture edge. Consequently, a Lyot stop, \( d \), can block all the on-axis light, leaving zero final field \( e \).
Example of band-limited function

\[
\tilde{M}(x) = N \left[ 1 - \frac{\sin(\pi \epsilon \lambda x)}{\pi \epsilon \lambda x} \right]
\]

\[
M(u) = N \left[ \delta \left(\frac{u}{D_\lambda}\right) - \frac{1}{\epsilon D_\lambda} \Pi \left(\frac{u}{\epsilon D_\lambda}\right) \right]
\]
Apodization
Apodized Pupil Lyot Coronagraph (Soummer et al.)

\[ \Phi_\text{undersize the Lyot stop.} \]

\[ \text{No. 1, 2009 APLCs FOR ARBITRARY APERTURES. II. 697} \]

\[ \text{transform is Hermitian and the eigenvalues} \]

\[ \text{the eigenfunctions are complex. In the following, we will only} \]

\[ \text{Basic layout of an APLC is similar to a classical Lyot coronagraph, with an upstream apodized pupil in plane A. Hard-edged focal masks are placed in the} \]

\[ \text{Ψ}_r(\text{sr e a la n ds y m m e t r i c}) \]

\[ \text{r = } \text{(1} \]

\[ \text{(Sivaramakrishnan & Lloyd} \]

\[ \text{approach is that there is only one physical apodizer involved, rejection factor of the starlight. An interesting aspect of this} \]

\[ \text{stages can be used. Each coronagraphic stage provides the same} \]

\[ \text{apodizing element is needed. In principle, several successive} \]

\[ \text{the second coronagraphic stage, the apodization is therefore} \]

\[ \text{field amplitude is proportional to the entrance pupil apodization,} \]

\[ \text{it is remarkable that these structures do not influence the field in the aperture} \]

\[ \text{both figures are displayed with appropriate scalings so that the apodizer inside} \]

\[ \text{it is remarkable that these structures do not influence the field in the aperture} \]

\[ \text{Figure 3.} \]

\[ \text{APLC with generalized prolate spheroidal apodization. As expected,} \]

\[ \text{asymptotically. APLCs take advantage of the particular shape of} \]

\[ \text{all the discontinuities in the pupil appear in negative, with no} \]

\[ \text{m.ore generally, prolate solutions} \]

\[ \text{basis over the aperture} \]

\[ \text{immediate consequence is that there are residual light is diffracted in the Lyot} \]

\[ \text{exist in any case, and we will study in Section} \]

\[ \text{this function, which saturates quickly, and thus enables very high} \]

\[ \text{as broadband, with} \]

\[ \text{monotonically decreasing with the} \]

\[ \text{method is also known in the field of signal processing and laser} \]

\[ \text{any part of achromatization is obtained on manufacture} \]

\[ \text{as interferometric methods can be used to produce} \]

\[ \text{In the case where the support structures are not included in the} \]

\[ \text{amplitude by the pupil function. In the case illustrated in Figure} \]

\[ \text{condition, identical inside the aperture in both planes to a scaling} \]

\[ \text{is monotonically decreasing with the} \]

\[ \text{very close to 1, for relatively small} \]

\[ \text{the Lyot stop amplitude is proportional} \]

\[ \text{integral of the apodizer, residual light is diffracted in the Lyot} \]

\[ \text{onto a prolate function. In broadband, with} \]

\[ \text{pupil apodization (left), since the apodizer is the eigenfunction of the system.} \]

\[ \text{is remarkable that these structures do not influence the field in the aperture} \]

\[ \text{it is remarkable that these structures do not influence the field in the aperture} \]

\[ \text{Figure 4.} \]
Shaped pupil zoo (Kasdin et al.)

<table>
<thead>
<tr>
<th>Mask</th>
<th>Ring</th>
<th>Barcode</th>
<th>Cross-barcode</th>
<th>Spiderweb</th>
<th>Starshape</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Ring Mask" /></td>
<td><img src="image2" alt="Ring PSF" /></td>
<td><img src="image3" alt="Barcode Mask" /></td>
<td><img src="image4" alt="Barcode PSF" /></td>
<td><img src="image5" alt="Spiderweb Mask" /></td>
<td><img src="image6" alt="Starshape Mask" /></td>
</tr>
<tr>
<td><img src="image7" alt="Cross-barcode Mask" /></td>
<td><img src="image8" alt="Cross-barcode PSF" /></td>
<td><img src="image9" alt="S-K Mask" /></td>
<td><img src="image10" alt="Early ripple designs" /></td>
<td><img src="image11" alt="ripple1 Mask" /></td>
<td><img src="image12" alt="ripple2 Mask" /></td>
</tr>
<tr>
<td><img src="image13" alt="S-K PSF" /></td>
<td><img src="image14" alt="Early ripple designs PSF" /></td>
<td><img src="image15" alt="ripple1 PSF" /></td>
<td><img src="image16" alt="ripple2 PSF" /></td>
<td><img src="image17" alt="ripple3 PSF" /></td>
<td></td>
</tr>
</tbody>
</table>

- **Shaped pupils**: $A(x,y)$ is zero-one valued (holes in masks)
  - **Advantages**:
    - Simple to manufacture
    - Inherently broadband
    - Minimally sensitive to aberrations
    - No off-axis degradation of PSF
  - **Disadvantages**:
    - Throughput (though roughly the same as 8th order Lyot coronagraph)
    - IWA (better IWA can be achieved through less discovery space or greater simplicity)
Using mirror to shape beams

Pupil Mapping for Apodization

Phase-Induced Amplitude Apodization (PIAA, Guyon et al. 2003)
Phase masks

- Replace opaque Lyot mask by a transparent phase-shifting mask

[Diagrams showing the operation of phase masks]

Roddier phase mask

Four-quadrant phase mask (Rouan et al. 2000)
Optical vortex
Nye & Berry 1974

• Singularity, point of zero intensity in an optical field

• Generated by a screw dislocation
The optical vortex formula

\[ e^{il\theta} \]

\( l \) is the vortex topological charge

= height of the dislocation
Propagation of an optical vortex
Pupil plane $\Rightarrow$ Image plane

Vortex field

Uniform field
Propagation of an optical vortex
Image plane => Pupil plane

Vortex field
Uniform field
Put your 3D glasses on

- on-axis vortex
- no vortex/off-axis
Optical vortex coronagraph
Mawet et al. 2005, Foo et al. 2005

• An image-plane optical vortex perfectly clears out the downstream pupil from any on-axis light

• Use a diaphragm (Lyot stop) => IDEAL CORONAGRAPH
Basic single-stage Lyot/VVC layouts

Telecentric

Non-Telecentric
The maths behind the vortex coronagraph

\[ A_{pup}(\rho, \psi, l_p) = FT\left[ \frac{2J_1(2\pi R_{tel}r)}{2\pi R_{tel}r} e^{i(l_p\omega-\pi/2)} \right](\rho, \psi) \]

\[ \approx \int_0^\infty J_1(2\pi R_{tel}r)J_{l_p}(2\pi \rho r)dr \]

*Hankel transform of \( J_1 \) of order equals to the topological charge of the vortex*

\[ H_l \left[ J_1(r) \right] \]
Hankel transform has amazing properties

Solutions for $H_l \left[ J_1(r) \right]$ are known
(Weber-Schafheitlin integral)

$$H_l \left[ J_1(r) \right] = 0$$

for $r < R_{tel}$
for $l = 2, 4, 6, 8, ...$
Engineering coronagraphs
Amplitude masks

Hole

Free standing

Microdot
Phase masks

DZPM

FQPM

CMC

Etching / coating
Hybrid masks (phase + amplitude)

Nickel mask has been vacuum-deposited on a fused silica substrate. Attenuation profile was built up in a number of passes with a computer-controlled moving slit. The same mechanism will be used to superimpose a dielectric phase layer in future work.

Comparison of the prescribed transmittance profile with the measured profile of the mask pictured at left. Desired profile is the red curve, the measured profile is the blue curve.

8.4.5.1 Baseline Design, Hybrid Lyot Coronagraph

HLC has a small focal plane mask with carefully optimized layers of nickel and dielectric which create a profile of intensity and phase transmission that stops the bulk of the starlight and sends the rest of the light toward a so-called Lyot stop in the pupil plane, as shown in Figure 8-3 and Figure 8-4. The Lyot stop blocks the remaining starlight but passes the light from the planet. The hybrid Lyot masks that have been developed to date are image-plane masks that appear as a fringe pattern of metal deposited on glass, with an additional (thus hybrid) layer of dielectric to compensate for residual phase errors.

8.4.5.2 Alternate Coronagraph: PIAA

This section describes the PIAA coronagraph option for Exo-C. The PIAA coronagraph uses pairs of aspheric mirrors to reshape the pupil plane intensity distribution, resulting in a Gaussian-like (prolate spheroidal) distribution which eliminates diffraction sidelobes. In a flight configuration, there would be two DMs located before the input PIAA optics and a second (inverse) set of PIAA optics after the focal-plane mask prior to the science camera.

Most of the work on PIAA has been carried out at NASA JPL's HCIT and the Ames Coronagraph Experiment (ACE) testbed at NASA Ames Research Center (ARC) and a vacuum chamber at Lockheed Martin ATC, as shown in Figure 8-3.
Geometrical Phase masks

HWP-FQPM

charge 2 vortex

charge 4 vortex

sub-λ grating

8OPM

LCP

photonics crystal

Nickel mask has been vacuum-deposited on a fused silica substrate. Attenuation profile was built up in a number of passes with a computer-controlled moving slit. The same mechanism will be used to superimpose a dielectric phase layer in future work.

Comparison of the prescribed transmittance profile with the measured profile of the mask pictured at left. Desired profile is the red curve, the measured profile is the blue curve.
Liquid Crystal Polymers (LCP)

- birefringent (resolution < 1 μm)
- malleable (liquid)
- harden into a solid state (UV cured)
- structure sealed
- vacuum/cryo compatible
Subwavelength gratings

- Grating pitch $< \lambda/n$

- Form birefringence can be engineered at the nano-scale

- Chromatic leakage can be minimized through optimization

- Monolithic structures

- Downscaling to near-IR and visible underway (European funded research)

L' band Charge 2 VVC made from diamond sub-$\lambda$ gratings (installed at VLT and LBT, see Delacroix et al. 2013 & Mawet et al. 2013)
Photonic crystals
Murakami, Nishikawa et al.

- Same physics as for sub-\(\lambda\) gratings
- Fabrication technology different
- Autocloning instead of deep etching
  - Accurate center
  - Charge 2 only so far
- Visible devices tested (Japanese collaboration)
Summary: the gamut of coronagraphy

Amplitude

Lyot coronagraph
APLC
Shaped pupil

Hybridization

Phase

Phase masks
PIAA

Image plane
Pupil plane
## Solutions for existing facilities

<table>
<thead>
<tr>
<th>Facility</th>
<th>Lyot/HLC</th>
<th>APLC</th>
<th>SP</th>
<th>4QPM/8OPM</th>
<th>(RA)VC</th>
<th>PIAA/CMC</th>
<th>APP</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3K</td>
<td>PHARO</td>
<td>P1640</td>
<td></td>
<td></td>
<td>(SDC)/PHARO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LMIRCAM</td>
<td></td>
<td>LMIRCAM</td>
</tr>
<tr>
<td>MagAO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clio2</td>
</tr>
<tr>
<td>SCExAO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GPI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPHERE</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>WFIRST-AFTA</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(x)</td>
</tr>
</tbody>
</table>
4 Pillars of High Contrast Imaging

- Adaptive optics
- Coronagraphy
- Differential imaging
- Post-processing

- Know your star (age, L, distance, proper motion, etc.)!
Do you see the planet after the coronagraph? No?

Contrast

Angular separation ($\lambda/D$)

1

$10^{-6}$

$1\lambda/D$

speckle noise

$0$ $5\lambda/D$

Angular separation ($\lambda/D$)
Red pill: image plane wavefront sensing

Blue pill: differential imaging

I’ll have both!
Let’s cheat: the planet is there
Rotate telescope (or track pupil with rotator) $45^\circ$

Contrast

Angular separation ($\lambda/D$)

$10^{-6}$

$0 \quad 5 \lambda/D$
Subtract, do you see the planet now?

Angular differential imaging (ADI)
ADI at small inner working angles (IWA)
Self subtraction

\[ \frac{\lambda}{D} \]

\[ 10^{-6} \]

\[ 1\lambda/D \]

\[ 1\lambda/D \]
Spectral differential imaging (SDI)

- Requires dual beam imagers or integral field spectrographs (P1640, GPI, SPHERE, CHARIS)
- Speckles scale as $\lambda$
- Real objects don’t move
- Suffers from self subtraction at small IWA too
Reference star differential imaging (RDI)

- Observe another similar star close in time, with as little telescope motion as possible
- Polarization differential imaging (PDI)
  - NO geometrical limitations at small IWA
Polarization differential imaging (PDI)

- Very efficient speckle calibration (starlight unpolarised, watch out for IP)
- Precious information about scatterer (composition, porosity, etc.)

\[ \text{NACO Ks band PDI (Milli, Mawet, Lagrange, et al. in preparation)} \]

\[ M^3: \text{Milli, Mouillet, Mawet et al. 2013} \]
4 PILLARS OF HIGH CONTRAST IMAGING

- Adaptive optics
- Coronagraphy
- Differential imaging
- Post-processing

- Know your star (age, L, distance, proper motion, etc.)!
Best way of combining data

Image/signal processing community:

Kernel PCA, Sparse PCA, Randomized PCA, Probabilistic PCA, ICA, Factor Analysis,
Sparse coding / dictionary Learning
Locally optimized combination of images (LOCI)

- LOCI = Locally Optimized Combination of Images ≠ LUCKY imaging
- Invented by C. Marois to reduce Angular Differential Imaging data.
- Refined by D. Lafreniere
- Formalism exposed in Lafreniere et al. 2007
- Applicable to ANY kind of speckle calibration (reference star subtraction, spectral differential imaging, polarization differential imaging, ADI).
Goal: find the optimal linear combination of reference images to reduce noise in the optimization section of the target image.
Principle

“By optimizing the weights given to the N available reference PSF images according to the residual noise obtained, this approach produces a representation of the target PSF image that is better than any predefined combination of the reference PSF images. Furthermore, it is advantageous to optimize the coefficients of the linear combination for subsections of the image because the correlation between the target and the reference PSF images generally varies with position within the target image”
Construct super-reference from the set of reference images, section by section

Reference image data cube

\[ C_i? \]

\[ \min \sigma^2 = \sum_{C_i} (T-SR)^2 = \sum (T-\sum c_i R_i)^2 \]

\[ \frac{\partial \sigma^2}{\partial c_i} = 0 \quad \Rightarrow \quad Ax = b \]

\[ A_{jk} = \sum R_j R_k \quad x = c_j \quad b_j = \sum R_j T \]

\[ SR = \sum c_i R_i \]
Why 2 sections? Size of the sections.

- One ultimately wants to construct the super reference pixel by pixel and know the optimal value (=optimal linear combination of reference image to subtract) to attribute to that individual pixel;

- If one applies LOCI to the pixel in question, it’ll just yield 0 in the final image after subtraction, because we simply ask him to do so!

- There is an optimal size for both sections: the optimization region must be big enough so that the signal from the companion is not subtracted out!

- In practice, 10-100 fwhm$^2$
Examples
HR8799’s 4 planets

Marois et al. (2008), Marois et al. (2010)
With NICMOS archival data

Soummer et al. (2011)
Principal Component Analysis

\[ \min \sigma^2 = \sum (T - SR)^2 = \sum (T - \sum <T,Z_i>Z_i)^2 \]

Karuhnen Loeve transform of the ensemble of references (or Principal Components over the covariance matrix): minimizes the least squared distance between any random PSF realization and the set of references.
Recipe (courtesy L. Pueyo, STScI)

1. partition the target \( T(n) \) and references \( R_k(n) \) images in an ensemble of \( \mathcal{S} \) areas, and subtract their average values so that they have zero mean over \( \mathcal{S} \).

2. compute the Karhunen-Loève transform of the \( \{R_k(n)\}_{k=1}^K \):

\[
Z_{KL}^k(n) = \frac{1}{\sqrt{\lambda_k}} \sum_{p=1}^{K} c_k(\psi_p) R_p(n),
\]

where the vectors \( C_k = [c_k(\psi_1), \ldots, c_k(\psi_K)] \) are the eigenvectors of the K-dimensional covariance matrix \( E_{RR} \) of the \( R_k(n) \) references, and \( \{\Lambda\}_{k=1}^K \) are its eigenvalues.

3. choose a number of modes \( K_{\text{clip}} \) to keep in the estimate \( \hat{l}_{\psi_0}(n) \).

4. compute the best estimate \( \hat{l}_{\psi_0}(n) \) of \( l_{\psi_0}(n) \) from the projection of \( T(n) \) on the \( \{Z_{KL}^k(n)\}_{k=1}^{K_{\text{clip}}} \) basis:

\[
\hat{l}_{\psi_0}(n) = \sum_{k=1}^{K_{\text{clip}}} <T, Z_{KL}^k>_{\mathcal{S}} Z_{KL}^k(n).
\]

5. calculate the final image \( F(n) = T(n) - \hat{l}_{\psi_0}(n) \),

\[
F(n) = \left( T(n) - \sum_{k=1}^{K_{\text{clip}}} <T(n), Z_{KL}^k>_{\mathcal{S}} Z_{KL}^k(n) \right)
\]

This is what SVD does.
4 Pillars of High Contrast Imaging

- Adaptive optics
- Coronagraphy
- Differential imaging
- Post-processing

- Know your star (age, L, distance, proper motion, etc.)!
Fig. 6.—The current census of companions in the brown dwarf (green) and planetary (blue) mass regimes that have both age and bolometric luminosity measurements from the compilation in Table 1. Many companions lie near the deuterium-burning limit while only a handful of objects are unambiguously in the planetary-mass regime. Hot-start evolutionary models are from Burrows et al. (1997); orange, green, and blue tracks denote masses $>80M_{\text{Jup}}$, 14–80$M_{\text{Jup}}$, and $<14M_{\text{Jup}}$.

- For young field stars, distant stellar companions can help age-date the entire system. For example, the age of Fomalhaut was recently revised to $\sim 400$ Myr from $\sim 200$ Myr in part due to constraints from its wide M dwarf companions (Mamajek 2012; Mamajek et al. 2013). Ultimately, if the age of a host star is unknown, the significance and interpretation of a faint companion is limited if basic physical properties like its mass are poorly constrained.

- Epoch of planet formation. Planets take time to form so they are not exactly coeval with their host stars. Their ages may span the stellar age to the stellar age minus $\sim 10$ Myr depending on the timescale for giant planets to assemble. Planets formed via cloud fragmentation or disk instability might be nearly coeval with their host star, but those formed by core accretion are expected to build mass over several Myr. While this difference is negligible at intermediate and old ages beyond a few tens of Myrs, it can have a large impact on the inferred masses of the youngest planets ($\ll 20$ Myr). For example, if the age of the young planetary-mass companion 2M1207–3932 b is assumed to be coeval with the TW Hyriade Association ($\tau = 10 \pm 3$ Myr) then its hot-start mass is $\approx 5M_{\text{Jup}}$. On the other hand, if its formation was delayed by 8 Myr ($\tau = 2$ Myr) then its mass is only $\approx 2.5M_{\text{Jup}}$.

- Atmospheric models. Atmospheric models can influence the inferred masses of imaged exoplanets in several ways. They act as surface boundary conditions for evolutionary models and regulate radiative cooling through molecular and continuum opacity sources. This in turn impacts the luminosity evolution of giant planets, albeit minimally because of the weak dependence on mean opacity ($L(t) \propto \kappa_0^{3.5}$; Burrows & Liebert 1993; Burrows et al. 2001). Even the unrealistic cases of permanently dusty and perpetually condensate-free photospheres do not dramatically affect the luminosity evolution of cooling models or mass determinations using age and bolometric luminosity (Baraffe et al. 2002; Saumon & Marley 2008), although more realistic (“hybrid”) models accounting for the evolution and dissipation of clouds at the L/T transition can influence the shape of cooling curves in slight but significant ways (Saumon & Marley 2008; Dupuy et al. 2015b). On the other hand, mass determinations in color-magnitude space are highly sensitive to atmospheric models and can result in changes of several tens of percent depending on the specific treatment of atmospheric condensates. Dust reddens spec...
When giant planets may still be assembling through core accretion, young ages of stars can be probed using extremely AO systems. Moreover, these techniques can help identify non-redundant aperture masking and provide contrast imaging of planets.

Contrast imaging can probe planetary masses that have not yet formed. In practice, the youngest T Tauri stars are important targets for direct imaging surveys highlighting the scientific context, strengths, and drawbacks of each method.

A contrast-limited regime means smaller physical separations can be studied. The luminosity bifurcation of young stars is attractive because stellar luminosities plateau on the main sequence, while planets and brown dwarfs continue to cool with age.

Age and distance are key parameters for planet detection. The properties of the host star, particularly its temperature and luminosity, are also important. The contrast between planets and their host stars is lower than at older ages because stellar luminosities plateau on the main sequence, whereas planets and brown dwarfs continue to cool with age.

Figures 2 and 3 show how planet mass sensitivity changes with age and distance. The panels display the e-folding mass sensitivity as a function of age and distance, with the blue panel showing the contrast between planets and their host stars. The orange panel shows the 50% sensitivity contour based on the median NICI contrast curve.
INFLUENCE OF STAR SAMPLE ON DIRECT IMAGING SENSITIVITY

Fig. 3.—Typical sensitivity maps for high-contrast imaging observations of T Tauri stars (5 Myr at 150 pc), young moving group members (30 Myr at 30 pc), and field stars (5 Gyr at 10 pc). Young moving group members are “Golidlocks targets”—not too old, not too distant. Black curves denote 10% and 90% contour levels assuming circular orbits, Condon hot-start evolutionary models (Baraffe et al. 2003), and the median NICI contrast curve from Biller et al. (2013). Gray and orange circles are RV- and directly imaged companions, respectively (see Figure 1).

Fig. 4.—The census of members and candidates of young moving groups. Prior to 2010 the M dwarf members were largely missing owing to their faintness and lack of parallax measurements from Hipparcos. Concerted efforts to find low-mass members over the past few years have filled in this population and generated a wealth of targets for dedicated direct imaging planet searches. Known groups. Historically, most groups themselves and new members of these groups were found with the aid of the Tycho Catalog and Hipparcos, which provided complete space velocities for bright stars together with ancillary information pointing to youth such as infrared excess from IRAS; X-ray emission from the Einstein or ROSAT space observatories; strong Hα emission; and/or LiIλ6708 absorption. As a result, most of the faint low-mass stars and brown dwarfs have been neglected. In recent years the population of “missing” low-mass stars and brown dwarfs in young moving groups has been increasingly uncovered as a result of large all-sky dedicated searches (Figure 4; Shkolnik et al. 2009; Lépine & Simon 2009; Schlieder et al. 2010a; Kiss et al. 2010; Rodriguez et al. 2011a; Schlieder et al. 2012a; Schlieder et al. 2012b; Shkolnik et al. 2012; Malo et al. 2013; Moor et al. 2013; Rodriguez et al. 2013; Malo et al. 2014; Gagné et al. 2014; Riedel et al. 2014; Kraus et al. 2014b; Gagné et al. 2015b; Gagné et al. 2015c; Binks et al. 2015). Parallaxes and radial velocities are generally not available for these otherwise anonymous objects, but by adopting the UVW kinematics of known groups, it is possible to invert the problem and predict a distance, radial velocity, and membership probability. Radial velocities are observationally cheaper to acquire en masse compared to parallaxes, so membership confirmation has typically been accomplished with high-resolution spectroscopy. The exceptions are for spectroscopic binaries, which require multiple epochs to measure a systemic velocity, and rapidly rotating stars with high projected rotational velocities ($v\sin i$), which produce large uncertainties in radial velocity measurements. The abundance of low-mass stars in the field means that some old interlopers will inevitably share similar space velocities with young moving groups. These must be distilled from bona fide membership lists on a case-by-case basis (Barenfeld et al. 2013; Wöllert et al. 2014; Janson et al. 2014; McCarthy & Wilhelm 2014; Bowler et al. 2015c).

The current census of directly imaged planets and companions near the deuterium-burning limit are listed in Table 1. Many of these host stars are members of young moving groups. β Pic, 51 Eri, and possibly TYC 9486-927-1 are members of the β Pic moving group (Zuckerman et al. 2001; Feigelson et al. 2006; Deacon et al. 2016). HR 8799 and possibly κ And are thought to be members of Columba (Zuckerman et al. 2011). 2M1207–3932 is in the TW Hydrae Association (Gizis 2002). GU Psc and 2M0122–2439 are likely members of the AB Dor moving group (Malo et al. 2013; Bowler B. 2016).
Signal detection theory in high contrast imaging
SDT

<table>
<thead>
<tr>
<th>Detection</th>
<th>H$_1$: signal present</th>
<th>H$_0$: signal absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positive</td>
<td></td>
<td>False Positive</td>
</tr>
<tr>
<td>False Negative</td>
<td></td>
<td>True Negative</td>
</tr>
</tbody>
</table>

TPF = sensitivity = power = TP/(TP+FN)  
FPF = 1-specificity = 1-CL = FP/(FP+TN)

$TPF = \frac{TP}{TP+FN} = \int_{\tau}^{+\infty} pr(x|H_1)dx$  
$FPF = \frac{FP}{TN+FP} = \int_{\tau}^{+\infty} pr(x|H_0)dx$
Ideal case

\[ FP_F = \int_{-\infty}^{+\infty} pr(x|H_0)dx = \int_{-\infty}^{+\infty} N(\mu, \sigma^2)dx \]

where

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

\[ \int_{\mu-c-\tau}^{+\infty} pr(x|H_1, \mu_c)dx = TPF \]
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>
<thead>
<tr>
<th>Statistics</th>
<th>No. Resolution Elements</th>
<th>Expected (d) ((\sigma))</th>
<th>(\langle d \rangle) ((\sigma))</th>
<th>Standard Deviation ((d)) ((\sigma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>10³</td>
<td>5.0</td>
<td>5.4</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>10⁴</td>
<td>5.33</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁵</td>
<td>5.06</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁶</td>
<td>5.06</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>MR10</td>
<td>10³</td>
<td>7.7</td>
<td>9.3</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>10⁴</td>
<td>8.20</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁵</td>
<td>8.02</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁶</td>
<td>8.02</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>MR1</td>
<td>10³</td>
<td>13.5</td>
<td>14.9</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>10⁴</td>
<td>13.95</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁵</td>
<td>13.44</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁶</td>
<td>13.48</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>MR01</td>
<td>10³</td>
<td>18.2</td>
<td>18.7</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>10⁴</td>
<td>17.2</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁵</td>
<td>17.13</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10⁶</td>
<td>17.18</td>
<td>0.41</td>
<td></td>
</tr>
</tbody>
</table>
What happens at small IWA?

![Graph showing contrast vs. angular separation with different line styles for 5σ contrast and FPF.]

- $\sigma$, here actually is the empirical standard deviation.
- The true STD of the underlying distribution is unknown!
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 \( r \)
- \( s_2 \) = 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
After correction

In this section, we discuss the consequences of small sample statistics, and a fake companion injected at the level of resolution elements at radius 1.5λ/D, 3.5*5σ. For calculating contrast under the null hypothesis: contrast curves, and so here we provide a simple recipe for calculating contrast under the null hypothesis:

\[ \text{Corrected contrast} = \text{Corrected FPF} \]

The case of one object with a single final image available as a standalone routine in most languages (e.g., "cvf" in IDL, but there are similar mechanics of a third-party pipeline. It could also occur if one does not have access to, or master the inner workings of a third-party pipeline. It could also occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely, and would only occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely.

When investigating the detection limits for a single observation, the contrast penalties. In this limiting case, detection limits would directly be prescribed by the t-distribution in order to preserve confidence levels. The detection is now much clearer than in the case where the contrast computation is allowed up, one is subject to direct hits from the limited number of samples available. In the following, we will overcome the limitations imposed by small samples, and the case where it is conducted on parts of, or all of the individual frames from the observing sequence.

In this section, we discuss the consequences of small sample statistics, and a fake companion injected at the level of resolution elements at radius 1.5λ/D, 3.5*5σ. For calculating contrast under the null hypothesis: contrast curves, and so here we provide a simple recipe for calculating contrast under the null hypothesis:

\[ \text{Corrected contrast} = \text{Corrected FPF} \]

The case of one object with a single final image available as a standalone routine in most languages (e.g., "cvf" in IDL, but there are similar mechanics of a third-party pipeline. It could also occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely, and would only occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely.

When investigating the detection limits for a single observation, the contrast penalties. In this limiting case, detection limits would directly be prescribed by the t-distribution in order to preserve confidence levels. The detection is now much clearer than in the case where the contrast computation is allowed up, one is subject to direct hits from the limited number of samples available. In the following, we will overcome the limitations imposed by small samples, and the case where it is conducted on parts of, or all of the individual frames from the observing sequence.

In this section, we discuss the consequences of small sample statistics, and a fake companion injected at the level of resolution elements at radius 1.5λ/D, 3.5*5σ. For calculating contrast under the null hypothesis: contrast curves, and so here we provide a simple recipe for calculating contrast under the null hypothesis:

\[ \text{Corrected contrast} = \text{Corrected FPF} \]

The case of one object with a single final image available as a standalone routine in most languages (e.g., "cvf" in IDL, but there are similar mechanics of a third-party pipeline. It could also occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely, and would only occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely.

When investigating the detection limits for a single observation, the contrast penalties. In this limiting case, detection limits would directly be prescribed by the t-distribution in order to preserve confidence levels. The detection is now much clearer than in the case where the contrast computation is allowed up, one is subject to direct hits from the limited number of samples available. In the following, we will overcome the limitations imposed by small samples, and the case where it is conducted on parts of, or all of the individual frames from the observing sequence.

In this section, we discuss the consequences of small sample statistics, and a fake companion injected at the level of resolution elements at radius 1.5λ/D, 3.5*5σ. For calculating contrast under the null hypothesis: contrast curves, and so here we provide a simple recipe for calculating contrast under the null hypothesis:

\[ \text{Corrected contrast} = \text{Corrected FPF} \]

The case of one object with a single final image available as a standalone routine in most languages (e.g., "cvf" in IDL, but there are similar mechanics of a third-party pipeline. It could also occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely, and would only occur if one does not have access to, or master the inner workings of a third-party pipeline. However, this situation is however unlikely.

When investigating the detection limits for a single observation, the contrast penalties. In this limiting case, detection limits would directly be prescribed by the t-distribution in order to preserve confidence levels. The detection is now much clearer than in the case where the contrast computation is allowed up, one is subject to direct hits from the limited number of samples available. In the following, we will overcome the limitations imposed by small samples, and the case where it is conducted on parts of, or all of the individual frames from the observing sequence.
Solution: more information about the noise properties

![Graph showing contrast penalty vs debinning factor](image)

- Photon noise!
- Ergodicity!
What if underlying noise is not Gaussian

![Graph showing FPF (1-CL) vs Angular separation (\(\lambda/D\)) for different scenarios.](image)