



# An Introduction to High Contrast Differential Imaging of Exoplanets and Disks

Katherine B. Follette

Amherst College, Department of Physics and Astronomy, USA

Received 2023 May 5; accepted 2023 July 27; published 2023 September 18

## Abstract

This tutorial is an introduction to High-Contrast Imaging, a technique that enables astronomers to isolate light from faint planets and/or circumstellar disks that would otherwise be lost amidst the light of their host stars. Although technically challenging, high-contrast imaging allows for *direct* characterization of the properties of circumstellar sources. The intent of the article is to provide newcomers to the field a general overview of the terminology, observational considerations, data reduction strategies, and analysis techniques high-contrast imagers employ to identify, vet, and characterize planet and disk candidates.

*Unified Astronomy Thesaurus concepts:* [Exoplanets \(498\)](#); [Circumstellar disks \(235\)](#); [Direct imaging \(387\)](#); [Coronagraphic imaging \(313\)](#)

## 1. Introduction

One of the breakthrough technologies of modern exoplanet astronomy is the technique of high-contrast imaging (HCI, often referred to more simply as “direct imaging”). HCI is a catchall term that encompasses the instrumental hardware, image processing techniques, and observing strategies that are employed to enable astronomers to image very faint sources (planets, circumstellar disks) in the vicinity of bright stars.

This article provides a basic introduction to the challenge of high contrast imaging in Section 1. It then defines and briefly describes the hardware involved in HCI in Section 2. In Section 3, it outlines how hardware and atmospheric aberration manifest in the anatomy of a HCI Point-Spread Function (PSF). Section 4 introduces the range of “differential imaging” observational techniques that are employed to facilitate separation of starlight from disk or planet light in post-processing, and Section 5 outlines the algorithms used to do so. Section 7 describe analysis techniques commonly employed to extract the properties of imaged planets and disks from post-processed HCI images, and Section 8 describes potential sources of false positives. Technologies that complement HCI are covered briefly in Section 9. The article is accompanied by a python code tutorial containing sample implementations of each of the main differential imaging techniques, as well as exercises for the reader. It is available at [https://github.com/kfollette/PASP\\_HCItutorial](https://github.com/kfollette/PASP_HCItutorial).

Throughout this article, I include definitions of many terms and phrases peculiar to High-Contrast imaging, but also assume knowledge of some common astronomy and optics terms that readers just getting started in the field may not yet be familiar with. The references I have chosen to include in the main text are primarily to the foundational work(s) that developed a particular technique. They are intended merely as a starting point, and should not be interpreted as the “state of the art” in the field. Two living documents accompany the tutorial and provide additional background in both areas. The first (available at <https://bit.ly/HCIjargon>) provides definitions of key astronomy and optics jargon used throughout this tutorial, which some readers may find useful when they encounter unfamiliar terms. The second (available at [bit.ly/beginHCI](https://bit.ly/beginHCI)) provides a recommended reading and viewing list for those who would like to delve deeper into the techniques discussed here.

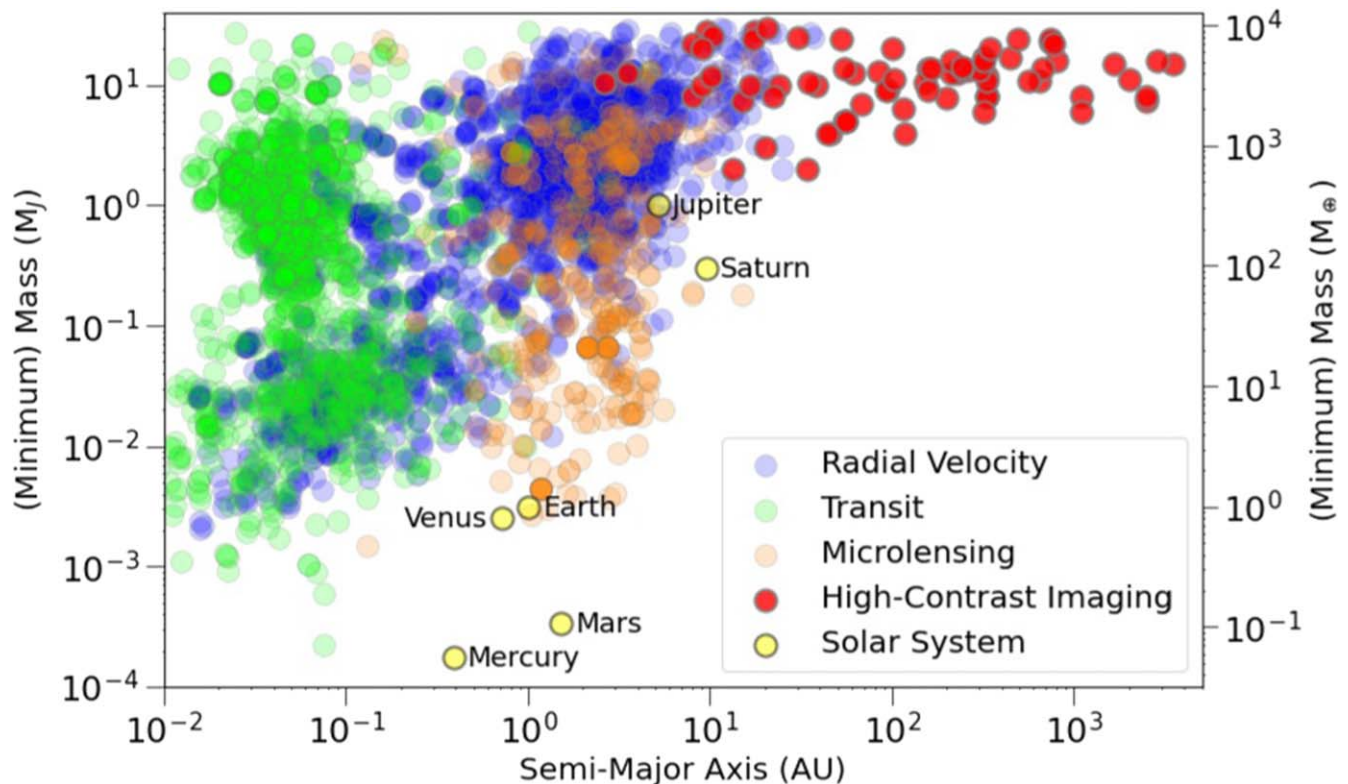
### 1.1. What is High-contrast Imaging?

The High-Contrast Imaging (HCI) technique is a relative newcomer in the world of exoplanet detection techniques, with the first discoveries in 2004 and 2008 (Chauvin et al. 2004; Kalas et al. 2008; Marois et al. 2008). Although the number of planet detections is currently lower for high-contrast imaging<sup>1</sup>

<sup>1</sup> As of the writing of this tutorial, ~50 companions have been imaged with estimated masses below the canonical “deuterium-burning” limit of  $<13M_J$  (the formal boundary between “planet” and “brown dwarf,” though the utility of this boundary as a defining line between populations is debated). However, this number more than doubles when considering all bound substellar ( $<70M_J$ ) companions to higher mass stars. Brown dwarf companions with masses less than  $\sim 20M_{Jup}$  are often referred to as “Planetary Mass Companions” (PMCs), and are likely part of the same underlying population as (i.e., formed similarly to) many of the objects currently classified as directly imaged “planets” (Wagner et al. 2019).



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



**Figure 1.** The population of known exoplanets discovered with high-contrast imaging (red) as compared to those found with indirect methods: transits (green), radial velocity (blue), and microlensing (orange) as of 2023 February per the [NASA Exoplanet Archive](#). Exoplanets are shown relative to solar system planets (yellow), highlighting the fact that detection techniques are not yet capable of detecting solar system analogs.

than for indirect (radial velocity, transit, and microlensing) techniques, directly imaged companions are arguably the best characterized exoplanets. HCI also provides the best prospects for current and future characterization of exoplanet atmospheres, particularly temperate ones conducive to life as we know it. The commitment of the community to HCI is evident in the first theme of *Pathways to Discovery in Astronomy and Astrophysics for the 2020s* (also known as the Astro2020 Decadal Survey)—“Pathways to Habitable Worlds.” It calls for a “step-by-step program to identify and characterize Earth-like extrasolar planets, with the ultimate goal of obtaining *imaging* and *spectroscopy* of potentially habitable worlds” (pg. 2, National Academy of Sciences Engineering & Medicine 2021, emphasis mine). The gap between the modern directly imaged planet population and Earth-analogs is large in both mass and semimajor axis space (see Figure 1). However, while indirect planet detection methods are currently more sensitive to terrestrial planets, the decadal survey goal of *imaging* and *spectroscopy* of exo-Earths cannot be achieved without direct detection.

Although the current state of the art in HCI is imaging of  $>1M_J$  planets at  $\sim$ tens of au separations, the future of the technique is bright (pun intended!), and vigorous ongoing

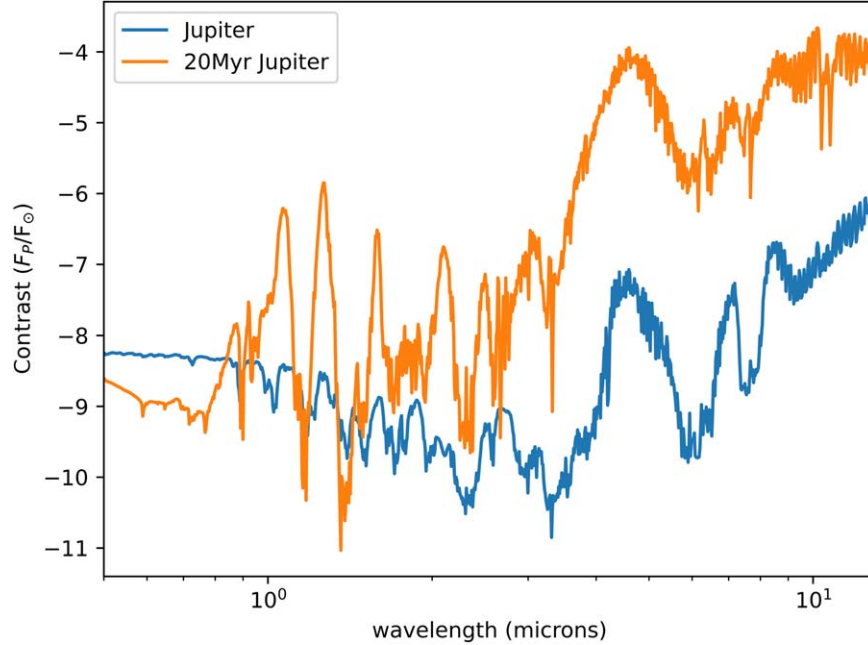
technology development will push its sensitivities to lower mass and more tightly separated planets.

### 1.2. What is Contrast?

In the context of HCI, the term “contrast” refers to the brightness ratio between an astronomical source (planet, disk) and the star it orbits. “High” contrast images are those where the ratio  $\frac{F_{\text{source}}}{F_{\text{star}}}$  is small, meaning the source is much fainter than the star—these detections are difficult. “Low” contrast images are therefore ones where the source-to-star ratio is larger, meaning the source is brighter relative to the star—these detections are less challenging.

Unlike stars, where absolute brightness is almost entirely a function of mass, for planets, brightness is a function of both mass and age. Planets begin their lives hot and bright and, lacking an internal source of energy sufficient to maintain that temperature, cool with time.

As they evolve, planetary spectra, and therefore contrast, change drastically. Figure 2 shows contrast at a range of wavelengths for the same planet (Jupiter) when “young” (20Myr) and “old” (4.5 Gyr, the age of our solar system). It



**Figure 2.** The predicted contrast ratios required to image Jupiter both as an “old” (4.5 Gyr, blue) and “young” (20 Myr, yellow) planet as a function of wavelength. Thermal and reflected light spectra were generated for both planets with PICASO (Batalha et al. 2019) and binned to a spectral resolution of 300. The young Jupiter’s spectrum was generated using the SONORA cloud-free atmospheric model grid (Marley et al. 2021) and divided by a simulated spectrum for a star with properties appropriate for the young Sun ( $T = 4300$  K,  $\log g = 4.3$ ,  $R = 1.2R_{\odot}$ ; Baraffe et al. 2015). The “old” Jupiter’s spectrum was generated for a 90% cloudy/10% cloud-free surface and divided by a solar spectrum.

highlights the extreme variation in contrast as a function of wavelength as planets age.

In thinking about contrast for point sources, it is useful to keep several benchmark quantities in mind, namely:

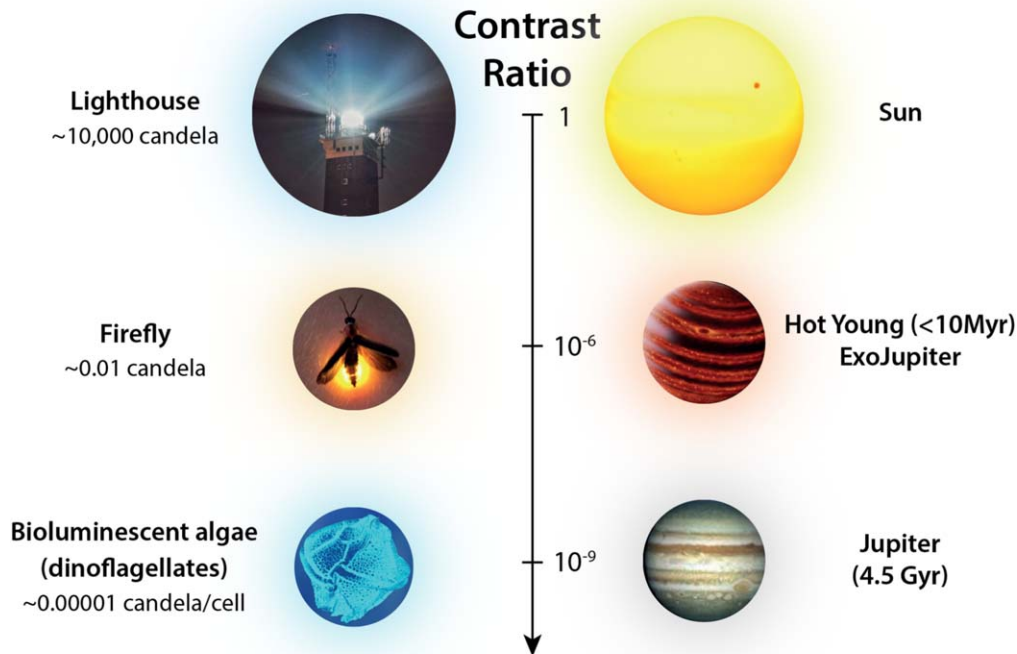
1. In the near-infrared ( $1\text{--}3\ \mu\text{m}$ ), young ( $\sim\text{few to few tens of Myr}$ ) giant planets generally have contrasts in the range  $\sim 10^{-5}\text{--}10^{-6}$  relative to their host stars. They radiate away much of their initial thermal energy over the course of the first tens of millions of years after formation, thus higher contrasts are required to detect them as they get older.
2. At  $3\text{--}5\ \mu\text{m}$ , the same young ( $\sim\text{few Myr}$ ) planets, have more moderate contrasts of  $\sim 10^{-3}\text{--}10^{-4}$ . With temperatures of  $\sim 500\text{--}1500$  K, this is because their thermal emission peaks in this wavelength regime, and the brightness gap relative to the much brighter and hotter (peak emission bluer) star is narrowed. This remains the region of most favorable contrast even as planets age.
3. In the optical, planets have undetectably low levels of direct thermal emission, and are seen instead in reflected light (stellar photons redirected/scattered by their atmospheres toward Earth). For mature planets ( $\gtrsim 100$  Myr), this wavelength regime provides more moderate contrasts than the NIR. For example, at 4.5 Gyr, Jupiter and Earth have contrasts of  $\sim 10^{-9}$  and  $10^{-10}$ , respectively at

$0.5\ \mu\text{m}$ . Combined with resolution advantages inherent in shorter wavelength imaging (See Section 2 for details) optical wavelengths provide the best prospects for future detection of solar system analog planets.

A simple analogy will help drive home the near (but not wholly) intractable nature of the contrast problem. As shown in Figure 3, for thermal emission from hot young exojupiters, the contrasts outlined above are comparable to the ratio of light emitted by a firefly relative to a lighthouse. For true (4.5 Gyr) Jupiter analogs in optical reflected light, a more apt comparison is a single bioluminescent alga relative to a lighthouse. This highlights the tremendous technological barriers that the field must overcome in order to achieve direct characterization of mature, potentially habitable exoplanets.

Precisely how hot a planet is at formation (and therefore how bright it appears) depends on how it was formed, and a range of formation modes are likely to overlap within the exoplanet population. In other words, planets (and brown dwarfs) of the same mass may have formed via different mechanisms, and their brightnesses therefore yield clues to their formation pathway.

Planets like those in our solar system most likely formed via a “cold start” mechanism involving the gradual assembly of solid material within a circumstellar disk. Their “cold” starts are only cold in comparison to so-called “hot start” planets, which also form in a circumstellar disk, but rapidly as a result



**Figure 3.** A schematic illustration of the magnitude of the brightness differential between the Sun and a hot, young exojupiter in the NIR and the Sun and a reflected light Jupiter in the optical. The brightness differential for a young Jupiter analog is  $\sim 10^{-6}$ , comparable to the brightness differential between a lighthouse and a firefly. Once a Jupiter-like planet has radiated most of the energy of formation and no longer glows brightly in the infrared, this differential drops to  $10^{-9}$ , akin to the brightness differential between a lighthouse and a single bioluminescent alga cell.

of gravitational collapse. The high masses and wide separations of most directly imaged planets make them good candidates for hot start formation, but current and next-generation instruments are detecting lower mass, closer-in planets for which formation mechanism is more ambiguous. The range of formation models and their predictions and assumptions is well-described in Spiegel & Burrows (2012). For our purposes, the most important takeaways are that directly imaged exoplanet brightnesses can only be translated to mass estimates under assumptions of: (a) stellar age, and (b) planetary formation pathway/initial entropy of the planet (unless a direct measure of the planet’s mass is available from another method, such as astrometry or radial velocity).

#### 1.2.1. What do We Learn from HCI Planet Detections?

The simplest measurements made for individual directly detected exoplanets are their **locations**<sup>2</sup> (astrometry) and **brightnesses** (photometry). Together with *evolutionary models* for young giant planets (which assume a formation pathway, e.g., Baraffe et al. 2003), photometric data allow for inference

of a planet’s *mass*, provided the system has a well-constrained **distance**<sup>3</sup> and a moderately constrained *age*.

Given the difficulty of robustly estimating ages for young objects, the preferred targets for direct imaging surveys have been young moving group stars; age estimates for these coeval groups are better constrained by averaging across independent estimates for their many members. Planetary luminosity and age can also be compared to the predictions of various planet formation models (e.g., the so-called cold/warm/hot start models, Spiegel & Burrows 2012) to inform the initial conditions under which planets are born.

Combining detection limits of large HCI planet-finding campaigns and evolutionary models allows for constraints on the *occurrence rates* of populations of exoplanets in various mass and separation ranges unique to direct imaging (currently  $\gtrsim 1M_J$  and  $\gtrsim 10$  au). Population constraints, in turn, inform formation models. For a review of what was learned about planet populations from the first generation of HCI campaigns, see Bowler (2016).

<sup>2</sup> In this section, I will place observed properties in **bold** the first time I reference them, and inferred physical properties in *bolded italics*.

<sup>3</sup> Nearly all HCI detections are for objects in the solar neighborhood, for which Gaia distances are sufficiently robust to consider them directly measured, rather than inferred quantities. For non-parallax distance measurements, this is not necessarily true.



Orbital monitoring of directly imaged planets also provides constraints on the dynamical evolution of young planetary systems. For example, *coplanarity* and the prevalence of *orbital resonances* in multi-planet systems inform planet formation and migration models (e.g., Konopacky & Barman 2019). *Alignment* (or misalignment) of planetary orbits with the stellar spin axis and/or the circumstellar disk plane informs the history of dynamical interactions within the system (e.g., Brandt et al. 2021; Balmer et al. 2022). Similarly, dynamical characterization of planets in systems with disk features hypothesized to be planet-induced provides a means to test disk-planet interaction models (e.g., Fehr et al. 2022). For a comprehensive review of planetary dynamical processes, see Davies et al. (2014) and Winn & Fabrycky (2015).

Finally, **spectroscopy** of imaged companions allows for direct characterization of *atmospheric properties*. To first order, low resolution spectra can inform the bulk *composition* of the atmosphere in more detail than photometry alone. For instance, even a low-resolution infrared spectrum of a giant planet can inform whether its atmosphere is CH<sub>4</sub> or CO-dominated. Directly imaged planet spectra, in combination with detailed atmospheric models, can also inform the *temperature-pressure structure* of the atmosphere, likely *condensate (cloud) species*, and even the prevalence of photo- and disequilibrium *chemical processes*. Constraints on *C/O ratios* of planetary atmospheres are probes of their formation locations relative to various ice lines that determine whether these elements are found in the gas or solid phase.

The advent of medium resolution spectroscopy of directly imaged planets with instruments such as VLT GRAVITY ( $R \sim 500$  in medium resolution mode) is enabling stronger constraints on these properties, with upgrades planned at the VLT to improve resolutions even further. Very high-resolution spectra of directly imaged companions will be enabled by coupling focal-plane optical fibers to existing high-resolution ( $R \sim 30,000$ ) spectrographs (e.g., The Keck Planet Imager and Characterizer (KPIC), Mawet et al. 2016). Such work requires very precise knowledge of planet astrometry to enable fiber placement, but will enable very exciting science such as constraints on planetary *rotation rates*, which can be compared against the predictions of various formation models. For a review of spectroscopy of directly imaged planets, see Biller & Bonnefoy (2018) and Marley et al. (2007).

### 1.2.2. What do We Learn from HCI Disk Detections?

HCI’s detection efficiency is significantly higher for circumstellar disk structures than for planets,<sup>4</sup> and many high-

resolution high-contrast images of circumstellar material have been collected by exoplanet direct imaging surveys (e.g., Avenhaus et al. 2018; Esposito et al. 2020; Rich et al. 2022). Such observations provide direct constraints on the distribution and composition of planet-forming material. Symmetric morphological features (such as rings, gaps, and cavities), inform the distribution of dust in planet-forming systems and, likely, the architectures of their planetary systems. Asymmetric features (such as warps and spiral arms) provide indirect evidence of embedded or undetected planetary perturbers or, perhaps, likely locations for future planet formation. Disk “signposts” of planet formation, though difficult to interpret, provide a wealth of information about planets and planet formation *at or near the epoch of formation*. For a comprehensive review of the state of high-contrast disk imaging, see Benisty et al. (2023).

NIR HCI disk images are also extremely powerful in combination with high-resolution millimeter imagery. In the millimeter and submillimeter, dust continuum emission traces large grains in the disk midplane, and millimeter line emission can be used to trace various gas-phase species as well. NIR high-contrast images trace an entirely different population, small micron-sized dust grains in the upper layers of the disk. Thus, the combination of NIR and mm high-resolution imagery yields a holistic picture of various disk components, a powerful combination for understanding the radial and vertical structure of disks.

Finally, multiwavelength NIR high-contrast imagery can be used to constrain grain properties such as size, porosity, and composition (e.g., Chen et al. 2020), as well as the water ice content of NIR-scattering grains (e.g., Betti et al. 2022). A good understanding of grain properties is essential to understanding the microphysics of dust coagulation, which will eventually form planets.

## 2. Enabling Technologies for High-Contrast Imaging

HCI is built upon a foundation of enabling technologies, namely: adaptive optics, coronagraphy, wave front sensing, and differential imaging techniques, each of which is introduced in this section. For a more comprehensive technical review of many of these technologies, see Guyon (2018).

### 2.1. Adaptive Optics

Adaptive optics is perhaps the most critical HCI enabling technology for ground-based imaging campaigns. Without it, image resolutions are limited by astronomical seeing, or the size of coherent patches in the earth’s atmosphere (approximated by the “Fried parameter”  $r_0$ , which has a  $\lambda^{6/5}$  dependence). With adaptive optics, modern HCI instruments

<sup>4</sup> The higher efficiency of disk detections is in part because of their ubiquity at detectable radii around young stars, but also because disks, unlike most planets, are detectable in polarized light. Polarized light from disks can be efficiently separated from unpolarized starlight via Polarized Differential Imaging (see Section 4.1).

**Table 1**

Seeing ( $r_0$ ) and Diffraction ( $\theta$ )-limited Resolutions at Three Common HCI Wavelengths for an 8 m Telescope at an Excellent Astronomical Site in Good Weather Conditions (0.''25 Seeing at V Band)

Distance	Resolution (in au)		
(pc)	@0.55 $\mu\text{m}$	@1.65 $\mu\text{m}$	@3.5 $\mu\text{m}$
Seeing-Limited Observations			
50	12.5	46.5	115
150	37.5	140	345
Diffraction-Limited Observations			
50	0.9	2.6	5.5
150	2.6	7.8	16.5

**Note.** Values are given in astronomical units for objects at distances of 50 pc (the volume limit of many HCI surveys) and 150pc (a typical distance to nearby star-forming regions).

can approach the diffraction limit,

$$\theta = 1.22 \frac{\lambda}{D}$$

where  $\lambda$  is the wavelength and  $D$  the diameter of the telescope. Table 1 gives the diffraction-limited resolution of an 8 m telescope at 0.55  $\mu\text{m}$  (V band), 1.6  $\mu\text{m}$  (H band) and 3.5  $\mu\text{m}$  (L band) in physical units as compared to the seeing limit at an exceptional telescope site under good weather conditions (0.''25 at 0.55  $\mu\text{m}$ ) at each wavelength.

The diffraction-limited Point-Spread Function (PSF)<sup>5</sup> of a circular telescope aperture is the so-called “Airy pattern.” In practical terms, this PSF places the majority of the incoming starlight into a “diffraction-limited core,” with a radius of  $1.22\lambda/D$  and a Full Width at Half Maximum (FWHM) of  $1.03\lambda/D$ . Extending from this central core are a characteristic set of “Airy” diffraction rings that decrease in amplitude outward and are spaced by roughly  $1\lambda/D$  from one another with the first minimum at  $1.22\lambda/D$ . In a perfect diffraction-limited system, the central “Airy disk” contains 84% of the total light in the PSF, with the remainder of the light in the Airy rings.

In the case of a telescope with a circular aperture and a central obscuration (e.g., by a telescope secondary mirror) the Airy pattern has a functional form of:

$$I(u) = \frac{1}{(1 - \epsilon^2)^2} \left[ \frac{2J_1(u)}{u} - \epsilon^2 \frac{2J_1(\epsilon u)}{\epsilon u} \right]^2$$

<sup>5</sup> A Point-Spread Function describes the appearance of a point source when imaged with a given combination of telescope, instrument, and wavelength. In functional form, it describes the location and intensity of light across the image plane.

where  $u$  is a dimensionless radial focal plane coordinate defined as:

$$u = \frac{\pi}{\lambda} D \theta$$

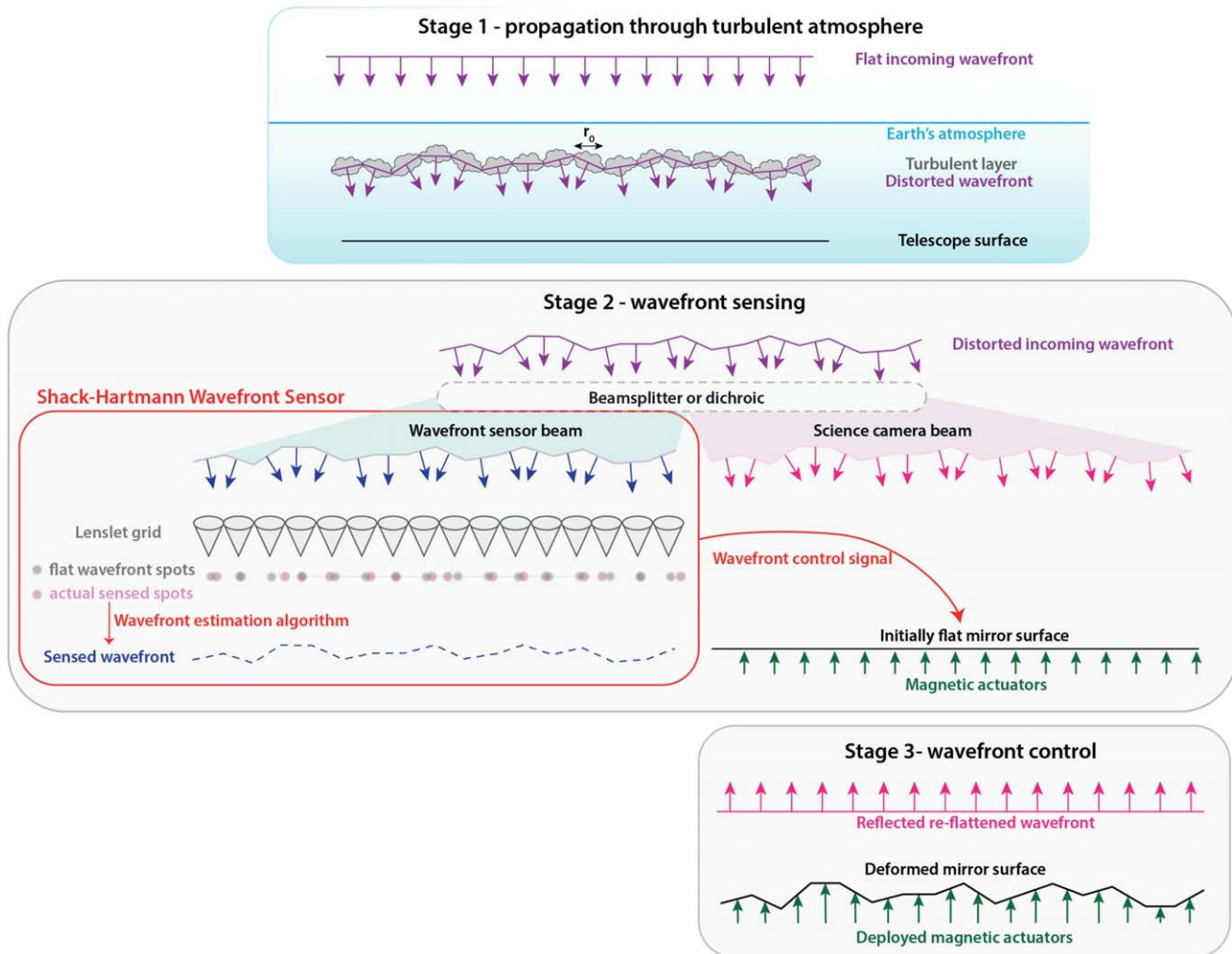
and  $\theta$  is the angle between the optical axis and the point of observation. The center of the PSF is at  $\theta = 0$  and therefore  $u = 0$ , and  $I(u)$  is the PSF intensity at location  $u$ . The quantity  $\epsilon$  is a measure of the amount of central obscuration expressed as a fraction of the total aperture (which acts to decrease the effective aperture and thus the predicted peak intensity), and  $J_1$  is the first order Bessel function of the first kind.

In practice, HCI PSFs tend to be dominated by Airy or Airy-like diffraction patterns with a few key deviations. First, no modern AO systems achieve perfectly diffraction-limited performance. The PSF of a modern adaptive optics PSF is often characterized by its so-called “Strehl Ratio” (SR), which is the ratio of a star’s observed peak intensity relative to that of its theoretical diffraction-limited peak intensity.<sup>6</sup> Modern Extreme Adaptive Optics (ExAO) systems routinely achieve SRs of 80%–95% in the Near Infrared, but only 10%–30% in the optical at present.

A proper treatment of the effect of the atmosphere on incoming starlight requires detailed atmospheric turbulence modeling (e.g., a Kolmogorov model). However, a decent first-order approximation of the effect of the Earth’s atmosphere on incoming starlight, depicted in Figure 4, is to imagine a plane-parallel electromagnetic wave<sup>7</sup> with some constant phase and amplitude encountering a layer in the Earth’s atmosphere composed of coherent patches of size  $r_0$  (atmospheric “cells”). Inside these cells, the wave front phase is aberrated such that it remains locally flat, however phase offsets occur between neighboring cells. Phase aberrations can take many forms and are often represented as an orthogonal basis set of polynomials with both radial and azimuthal dependencies (e.g., the Zernike polynomials). Low order aberrations have familiar names, and

<sup>6</sup> This theoretical PSF is not fully approximated by the relatively simple  $I_{u=0}$  described above for a given telescope aperture size ( $D$ ), central obscuration  $\epsilon$ , and wavelength  $\lambda$ , because (a) it assumes no other obscurations in the aperture (e.g., secondary mirror supports, downstream optical elements), and (b) it computes the PSF for a single wavelength, which is not measurable in practice. Thus, real Strehl Ratio approximations require detailed instrumental PSF models that include all of the telescope and instrument system’s optical elements. The on-sky predicted PSFs are then normalized to the same total intensity and divided to approximate Strehl Ratio. For further discussion of the subtleties of Strehl Ratio determination, see Roberts et al. (2004).

<sup>7</sup> Plane-parallel here means that if we were to draw a shape connecting equivalent phases of incoming electromagnetic waves from the same source, say the location where their electric field strengths are strongest, the shape of our equal-phase surface would be a plane perpendicular to the direction of travel. In other words, light from a distant source enters the Earth’s upper atmosphere in phase with neighboring light waves. This is an approximation because light exits a spherical object symmetrically, meaning that a surface of constant phase should always have some curvature; however, the distances to astronomical objects are vast compared to the sizes of the telescopes we use to intercept their light. This means that we intercept only a tiny area of a vast spherical shell of light from the star, a shell so vast that the tiny area we intercept can be treated as locally “flat.”



**Figure 4.** A simplified, schematic illustration of the process of adaptive optics. “Stage 1” depicts the effect of the Earth’s atmosphere on incoming plane-parallel light. The wave front is aberrated inside of locally coherent patches in the atmosphere, and enters the telescope aperture with corrugations of a characteristic size ( $r_0$ ). In “Stage 2”, the incoming light is passed through a beamsplitter or dichroic, which splits it, sending some to a wave front sensor and the rest to a science camera. In this case, a Shack–Hartmann wave front sensor (see Section 2.2) is depicted an array of lenslets in the focal plane. Each makes a spot whose location relative to the orientation of the lenslet is indicative of the slope of the incoming wave front. The spot locations are converted to a “best guess” of the incoming wave front shape and a corresponding control signal is sent to actuators under an (initially flat, generally tertiary) mirror. “Stage 3” depicts the result of the deformed wave front reflecting off of the deformed mirror, causing the reflected wave front to be re-“flattened,” thus compensating for atmospheric aberration. The sensed wave front is depicted here as an unrealistically perfect match to the true incoming wave front. In reality, kHz-scale time variation in the incoming wave front, unsensed or imperfectly estimated wave front aberration, and the speed and nature of the control algorithm mean that no wave front is perfectly sensed and corrected. Some residual corrugation will remain in any real AO system.

ones that you are likely to encounter in your annual eye exam, such as “astigmatism” and “coma.” Higher order aberrations take more complex forms in phase space, but all are essentially disruptions in the intrinsic shape of the incoming PSF. For illustrative purposes, let us imagine only the simplest two low-order modes, the so-called “tip” and “tilt” modes, which preserve the shape of the PSF but modify the direction of the incoming wave front relative to the original travel direction.

The effect of tip/tilt aberrations is that a wave front exiting a layer of atmospheric cells is no longer plane-parallel. Instead, it

is corrugated (the angle of arrival varies across the telescope aperture, see Figure 4’s “distorted incoming wave front”) with some wavelength-dependent characteristic length scale (The Fried coherence length,  $r_0 \sim \lambda^{6/5}$ ). For an atmospheric layer at a certain height, this characteristic length scale can also be represented as an angular scale called the “isoplanatic angle,”  $\theta_0$ . Note again that this is just a first-order approximation, albeit a useful one for building intuition, and that, in reality, there are a number of aberrating layers in the atmosphere with their own characteristic coherence lengths,

heights, and wind speeds. The practical consequence when integrated over the telescope aperture is that the light of each coherent patch manifests as its own diffraction limited PSF at a different location in the image plane centered around the optical axis of the telescope. The instantaneous result is a number of superposed independent images of the star equal to the number of coherent atmospheric patches that the wave front incident on the telescope passed through—i.e., the image is blurry.

Locally coherent patches at a given layer in the atmosphere only remain so on timescales of hundredths- to thousandths- of a second (due to wind, temperature/pressure variation, etc.) which means Adaptive Optics systems must operate on these timescales in order to detect and correct wave front aberrations with Wave front Sensors (WFS). Extending our toy example of an incoming plane-parallel wave front that experiences pure tip/tilt aberrations at a single layer in the atmosphere, imagine a series of corrugated wavefronts exiting this layer and being collected continuously by an astronomical detector over a realistic exposure time of several to several tens of seconds. The result will be a superposition of many hundreds or thousands of diffraction-limited PSFs (so-called “speckles”) at various locations relative to the central optical axis. The result is a seeing-limited PSF, whose size/FWHM will vary according to various properties of the atmosphere, but will always be much larger than the diffraction limit. Modern AO systems are able to operate at 1–2 kHz frequencies, however they are not able to perfectly sense the wave front, nor to perfectly or completely correct it on the relevant timescales. Many advancements are being made in both the hardware and software of wave front control, including the advent of algorithms that attempt to account for the time delay between sensing and applying a wave front correction by predicting the state of the wave front into the future (so-called “predictive control” algorithms, e.g., Poyneer et al. 2007; Guyon & Males 2017).

The consequence of a perfect AO system that could fully detect for and correct wave front aberration would be a perfect  $SR = 100\%$  diffraction-limited PSF. The reality is, of course, not perfect. A partially or imperfectly corrected wave front results in the alignment of many but not all of these instantaneous PSFs. Some uncorrected, residual seeing-limited “halo” with a width of approximately  $\frac{\lambda}{r_0}$  is expected, and its amplitude should decrease as the performance of the AO system (Strehl Ratio) improves. Imperfect wave front correction can also lead to certain persistent speckles, so-called “quasi-static speckles,” that are stable on timescales of minutes to hours. These are particularly worrisome because they can mimic planets, but they have the advantage of being static in their location in the instrument frame. They also exhibit spectra that are identical to that of the central star. These properties

make them amenable to removal by angular and spectral differential imaging (ADI/SDI, see Section 4).

NIR HCIs can have a dozen or more clear, detectable Airy rings in their unocculted AO PSFs. These Airy rings present a fundamental barrier to achieving high contrast in the environs of the central star, and additional optics are often employed to mitigate them. Because Airy rings are a consequence of diffraction at the edges of the entrance pupil, mitigating optics are generally pupil plane<sup>8</sup> optics that block light near its edges. One example is the “Lyot stop.”

The Airy PSF is also predicated on the assumption of a circular entrance aperture, which no realistic telescope entrance pupil is able to achieve. The presence of various optics, especially the secondary mirror and its supports, induce deviations from a perfect Airy PSF. To simulate an HCI PSF, therefore, requires a model of the telescope entrance aperture and any additional optics in the telescope beam. An example PSF for the Gemini Planet Imager is provided in Figure 5.

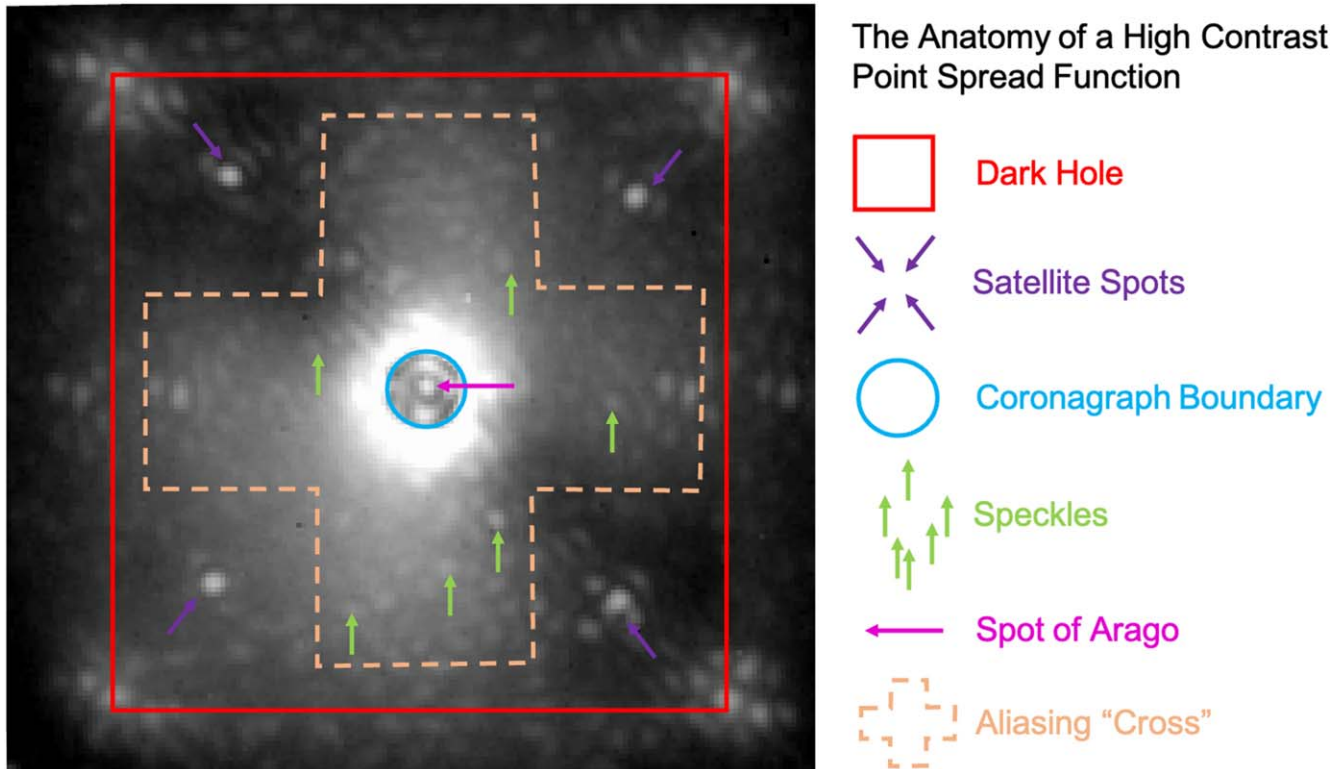
## 2.2. Wave front Sensing and Control

In addition to deformable mirrors (DMs), adaptive optics systems require instrumentation that can sense atmospheric aberrations and convert them to DM control signals on kHz frequencies. From an observer’s perspective, the most important features of this “Wave front Sensor” (WFS) and its accompanying control algorithm are its: wavelength, limiting magnitude, stability, gain, and cadence, each of which is described below.

*WFS wavelength*—WFS operate most often at optical wavelengths. Since most HCI is done in the NIR, such systems implement a dichroic that sends all optical light to the wave front sensor and all NIR light to the science camera. Although this results in no loss of light at the science wavelength, it does introduce a difference in the scale of the wave front aberrations that are sensed versus detected (namely,  $(\lambda_{\text{sensed}}/\lambda_{\text{detected}})^{6/5}$ ). NIR wave front sensing is an active area of development in HCI instrumentation for this reason. For a visible light HCI instrument, wave front sensing in the optical generally requires a beamsplitter, resulting in a substantial loss of signal to the science camera (50% or more) as light at the science wavelength is diverted to the WFS.

<sup>8</sup> Complex modern instruments utilize optics in both the “image plane,” where light incident on the telescope is brought to a focus, and the “pupil plane,” where light is collimated. Estimates of the appearance of an object in a given plane can be accomplished by Fourier transform of its appearance in the other. While image plane images show the on-sky source (often manipulated by upstream optics such as coronagraphs), pupil plane images are essentially images of the entrance aperture (a symmetry you can prove to yourself with a simple ray-tracing diagram), containing e.g., the central obscuration from the secondary mirror, the spider arms suspending it, etc. HCI instruments, especially those that require precise placement of optical elements in the pupil plane, are often equipped with “pupil-viewing” cameras, which image this entrance aperture.





**Figure 5.** A raw high-contrast image from the Gemini Planet Imager, with various features labeled. GPI’s square-shaped “dark hole” (region of AO correction) is marked in red. Satellite Spots injected intentionally into the images by the apodizer are shown with purple arrows, and serve as photometric and astrometric references. The central star is obscured by the coronagraph, the edge of which is depicted in blue. Diffraction does introduce some light to the region “underneath” the coronagraphic mask, including the “Spot of Arago” at the center of the image, marked in magenta. Examples of speckles, which are distributed throughout the image but are concentrated near the edge of the coronagraphic mask, are marked in green. Individual high-contrast imagers have various unique features, such as GPI’s “aliasing cross” (an optical effect caused by undersampling, see Poyneer et al. 2016).

*WFS limiting magnitude*—is a measure of the faintest targets for which the wave front can be effectively sensed, and is determined at the most basic level by the architecture of the WFS. Though there are many types of WFS, the most common are the Shack–Hartmann and Pyramid WFS. Tradeoffs in WFS qualities, such as sensitivity to wave front errors of various scales and linearity between WFS measurements and DM commands, determine the choice of WFS architecture (for a full discussion, see Guyon 2018). From the perspective of the observer, one practical consequence of WFS architecture is the range of magnitudes for which AO correction can be accomplished. A Shack–Hartmann WFS (SHWFS, see Figure 4 for a simple depiction) relies on a grid of lenslets placed in the pupil plane, each of which creates a spot on the WFS camera. The location of the spot created by each lenslet is controlled by the direction of the incoming wave front, and this shape can then be applied to the DM to correct aberrations. The limiting magnitude of a SHWFS is a fixed quantity determined by the required brightness for an individual lenslet spot to be sensed. Because the lenslets are physical optics, this cannot be modified without swapping out the grid of lenslets. A pyramid

WFS, on the other hand, modulates the incoming light beam around the tip of a four-faced glass pyramid, each facet of which creates an image of the telescope pupil on a WFS camera. These four pupil images can be analyzed to reconstruct the incoming wave front. A pyramid WFS camera’s pixels can also be binned to achieve correction on fainter guide stars. Although wave front information is lost in the binning process and the quality of the AO correction is therefore necessarily compromised, this does preserve the ability to apply (more modest) AO correction when imaging fainter stars.

*WFS stability*—is effectively a measure of how long and under what conditions a WFS can provide continuous adaptive optics correction. When AO systems are operating in “closed loop” mode, meaning corrections are being applied in real time, the loop will “open” in order to protect the DM if the sensed wave front deformations require corrections whose amplitudes are too great for the control range of the DM. This is called a “breaking” of the AO control loop. One of the more critical aspects of a wave front control algorithm is the “gain” applied to each sensed aberration.

*WFS control gain*—can be thought of as a multiplicative factor applied to the sensed wave front so that all of the sensed aberration is not corrected for at once, but instead some proportion of it. This is to avoid overcorrecting an aberration and driving the mirror into an oscillation, and also to allow unsensed or incorrectly sensed aberrations to pass without breaking the loop. Different sensed wave front aberrations (e.g., “low order” and “high order” modes) can have different gains, and this is one of the principal quantities that can be adjusted in real time during AO observations. Gain, wave front stability, WFS signal strength, and the nature of the control algorithm all conspire to determine the stability of the AO loop—basically its ability to remain closed during an observing sequence.

*WFS cadence*—is the timescale on which the wave front is sensed, and is the final factor controlling the quality and stability of AO correction. In this case, the wavelength of observation and nature of the telescope site (seeing, wind speed, etc.) sets the timescale on which the incoming wavefronts change, and the AO system must run faster than this timescale in order to apply high-quality correction. Many current AO systems operate at 1–2 kHz frequencies, with faster speeds being required at shorter wavelengths.

### 2.3. Coronagraphy

Another enabling HCI technology is coronagraphy, which utilizes one or more physical optics inside the instrument to suppress both direct and diffracted starlight before it reaches the detector. This allows for deeper imaging of planetary systems, as longer integration times can be used before saturation of the primary star. Coronagraphy is distinct from external occulters (“starshades”) and software algorithms (“wave front control”) that are designed to do similar things i.e. suppress and control light from the central star so that faint objects in its environs can be sensed. Available coronagraphic architectures have been rapidly expanding in recent years, and I will not provide a comprehensive review here, but will instead focus on the practical effects of a coronagraph for image processing.

The purpose of a coronagraph is to redirect starlight away from the image plane by blocking or modulating it with optical components, thus reducing the amount of light that must later be removed in post-processing in order to image faint companions.

Coronagraph optical components can modulate wave front amplitude or phase or, in many cases, both. The most basic coronagraphic architecture is an opaque or reflecting image plane spot in the center of the field, which prevents on-axis starlight from reaching the detector. Other coronagraphic architectures utilize interferometric techniques (e.g., the “vortex” coronagraph) to accomplish the same goal. Additional optics are often placed in the pupil plane to mitigate diffraction

around coronagraph edges and around the edges of the entrance aperture more generally, which effectively decreases the amplitude of the Airy rings and allows for higher contrast imaging and detection of fainter circumstellar signals.

## 3. The Anatomy of a High Contrast Image

Unlike many other fields of astronomy, raw HCI images rarely contain readily apparent raw signal from the circumstellar sources being targeted, even under aggressive hardware suppression of the stellar PSF. Post-processing is generally required to achieve the required contrast, and is covered in detail in Section 5. Nevertheless, the anatomy of a raw high-contrast image is important to understand in order to develop intuition for the range of artifacts that might survive into post-processing so they can be recognized and removed. This section lays out the anatomy of a “typical” coronagraphic high-contrast PSF, beginning with features at the center of the image and moving outward.

*Coronagraph and Spot of Arago*—the presence of a coronagraph in the beam results in a relative dearth of light at the center of the image. The angular size of the coronagraphic mask can be discerned in raw images by the ring of bright diffracted starlight just beyond its outer edge. Inside of this ring, the image is markedly darker, yet there is often a single brighter spot at the center, the so-called “spot of Arago” or “Poisson spot,” an artifact of Fresnel diffraction. This spot is not sufficiently bright to be used as a photometric or astrometric point of reference, however its detection and interpretation was central to our understanding of the wave nature of light, and it thus has a very important role in the history of optics.

*Optical Aberrations*—The evolving atmosphere and many optical elements of a high-contrast imaging instrument inevitably induce deviations in the PSF from the theoretical Airy Pattern. Many of these aberrations can be sensed and corrected by the Adaptive Optics system, but imperfectly, so some will survive into the final PSF, and cause its shape to deviate from an Airy pattern.

*Speckles*—The residual, uncorrected starlight that dominates raw high-contrast images generally comes in two forms. First, atmospheric or instrumental aberrations undetected or not fully corrected by the adaptive optics system manifest as “speckles” (images, often aberrated, of the central star) at a range of locations in the PSF, but concentrated toward the central optical axis. These evolve with the rapidly changing atmosphere, and blend into a diffuse halo of uncorrected starlight in most raw images (the so-called “seeing halo”). For very short exposures, speckles can be individually distinguished more readily, but in such cases they evolve quickly among images and thus rarely masquerade as planets in final PSF subtracted images. So-called “quasi-static” speckles are likely created by optical aberrations in the instrument and evolve much more

slowly, thus appear stably across multiple images and are more problematic. Various forms of active control have been developed to remove these quasi-static speckles (e.g., “speckle nulling,” Bordé & Traub 2006; Martinache et al. 2014) and many differential imaging processing techniques are designed specifically to distinguish quasi-static speckles from planets (see Section 4).

*Dark Hole/Control Region*—AO-corrected images also exhibit a boundary between the region of sensed wave front aberration/AO correction and an uncorrected/unsensed region. This boundary defines the so-called “dark hole” or “control radius” of an AO system. The location of this boundary in the image plane is a direct consequence of the wave front sensor’s inability to perfectly sense all pupil plane wave front aberrations. For example, there is a minimum size of wave front aberrations that an AO system can detect and correct, set by the spacing of actuators, wave front sensor optical component spacings (e.g., Shack Hartmann WFS), and/or wave front sensing camera pixel scales (e.g., for a Pyramid WFS). Any spatial frequency smaller than this limit cannot be corrected by the AO system, and this pupil plane limit maps to a particular location in the image plane. Thus, the image reverts to seeing limited outside of the boundary of the dark-hole, resulting generally in an increase in the intensity of the seeing halo at its boundary.

*Wind Artifacts*—Wind, particularly high altitude wind, drastically affects the speed at which the incoming wave front changes. AO systems therefore have a harder time “keeping up” with aberrations along one axis of the PSF (the wind direction) than others, and the AO correction is poorer along this axis. In most modern HCI imagery, the wind direction can be inferred from an apparent elongation of the speckle pattern in the wind direction (i.e., there are more speckles in the halo along the wind direction, where the AO system is struggling to “keep up”). This additional uncorrected light introduces a difference in the achievable contrast in an image azimuthally, with planets/disks that align with wind artifacts more difficult to detect.

*Satellite Spots*—One practical consequence of coronagraphy is the loss of a direct measurement of the central star’s astrometry and photometry. At the same time, photometric and astrometric characterization of substellar sources is dependent on measurement of these properties for the central star. For this reason, many modern HCI instruments inject reference “satellite” spots into images at known locations and with known brightness ratios relative to the central star. This is done either through a pupil plane optic custom-designed to inject satellite spots at certain locations and brightnesses or by using the deformable mirror of the telescope to produce them. Once photometrically and astrometrically characterized (e.g., Wang et al. 2014), these spots are sufficiently stable to serve as proxies for direct measurements of the location and brightness of the central star.

*Instrument throughput*—is a measure of the fraction of light entering the telescope aperture that ultimately makes it onto the detector. It is determined in part by the number of reflecting and refracting elements in the optical path, each of which results in loss of a few percent of incoming light. The operating wavelength of the science camera and wave front sensor is also a consideration. Generally wave front sensors have operated at shorter, visible wavelengths and HCI cameras have operated in the NIR, enabling a dichroic to be used to separate incoming light and minimize loss of light at the science wavelength. The advent of Infrared wave front sensors and visible light adaptive optics systems complicates this somewhat, to the extent that it can no longer be assumed that all light at the science wavelength is directed to the science camera. However, though clever combinations of filters and beamsplitters, as well as usage of light that is otherwise discarded by the system (e.g., by the coronagraphic occulter), help to maximize throughput in these cases.

## 4. Differential Imaging Techniques

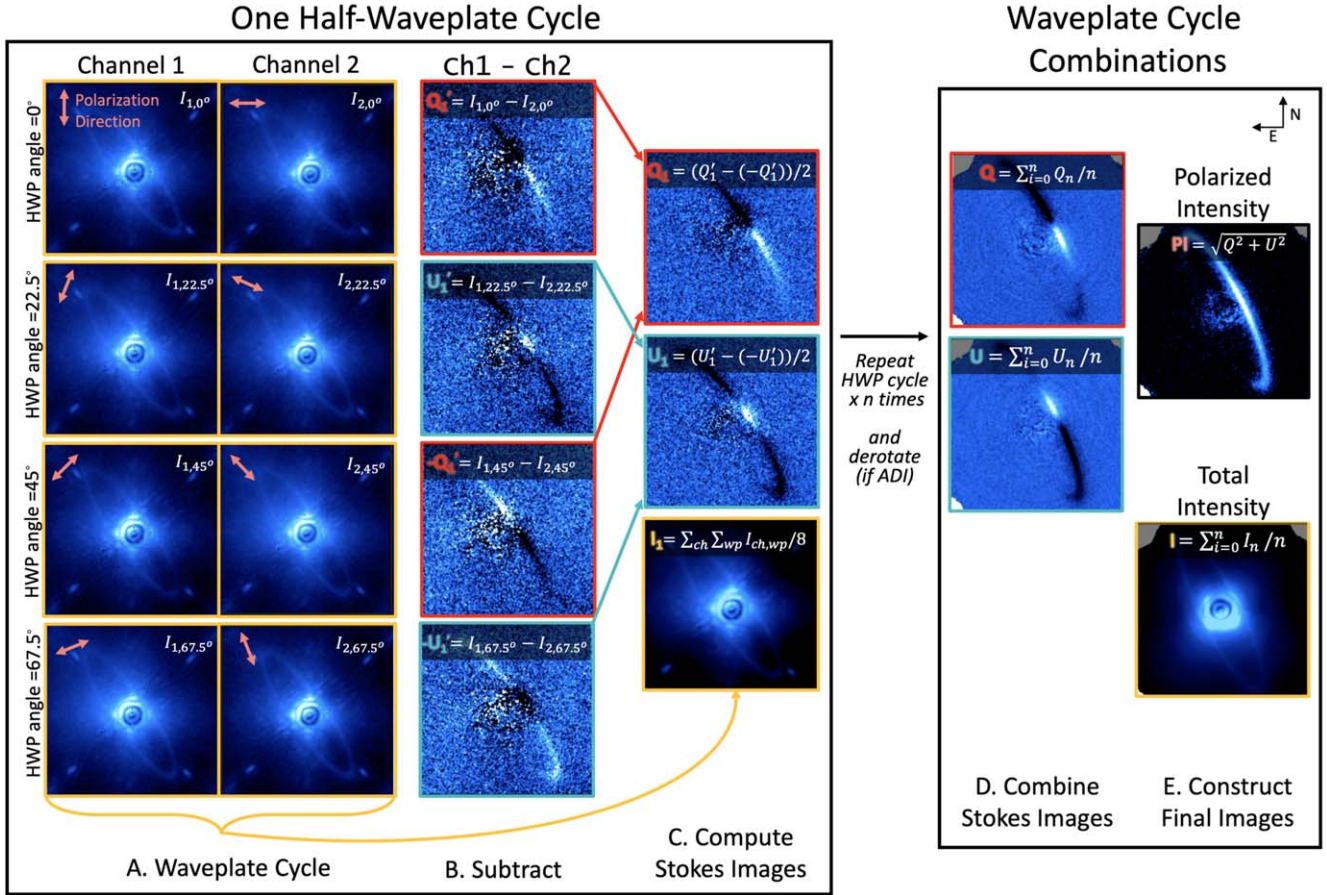
Ultimately, even the best HCI hardware can only suppress starlight by 3 or 4 orders of magnitude in brightness, still 2–3 orders of magnitude too low in contrast to image a hot young exo-Jupiter. Modern high-contrast imaging instruments rely on a number of clever data collection methodologies—collectively referred to as “differential imaging”—to facilitate separation of starlight from planet/disk signal. When distilled to their essence, all differential imaging techniques are designed to leverage wavelengths, angular locations, other sources, or polarization states where companion light is faint or absent to estimate and subtract the PSF of the central star. These techniques are presented here in rough order of “aggressiveness” in estimating and removing the PSF of the central star.

### 4.1. Polarized Differential Imaging (PDI)

Polarized Differential Imaging is the most common and successful technique for imaging circumstellar disk material in scattered light, and is shown schematically in Figure 6. PDI relies on the fact that light emitted directly from the central star is (generally) unpolarized. Dust grains in the circumstellar environment, on the other hand, preferentially scatter starlight with a particular polarization geometry. Scattering is most efficient for light with an electric field vector aligned orthogonal to both: (a) the line of sight from the scattering location on the disk surface to earth and (b) the vector connecting the scattering location and the central star. In principle, this means that disk scattered light signals should dominate PDI images, and (unpolarized) stellar emission should be absent.

PDI imaging separates incoming starlight according to the orientation of its electric field vector (i.e., it is linear polarization). An optic called a Wollaston prism accomplishes this by passing incoming light through a material that has





**Figure 6.** A schematic representation of the Polarized Differential Imaging (PDI) technique. Light from a disk-bearing star (in this case images of the debris disk host HR4796 A collected with the Gemini Planet Imager at K band) is split into two orthogonal polarization states (indicated in coral in the figure), and these two “Channels” (Column A’s “Channel 1” and “Channel 2”) are imaged simultaneously. A rotating Half-Wave plate (HWP) modulates the direction of both polarization directions by rotating 22°5 between images, in sets of four, at orientations of 0°, 22°5, 45°, and 67°5. The two simultaneously obtained orthogonal polarization channels are subtracted from one another (Column B). Subtractions for half-wave plate orientations 0° and 45° probe the Stokes  $Q$  parameter and its reverse. Subtractions for half-wave plate orientations 22°5 and 67°5 probe the Stokes  $U$  parameter and its reverse. These independent probes of Stokes  $Q$  and  $U$  can be combined (Column C) to average over location specific artifacts of the detector. The dual channels of Column A can also be combined across all 4 wave plate orientations to yield a Stokes  $I$  (total intensity) image (Column C, bottom). The cycle of 4 wave plate orientations is repeated a number of times, often with Angular Differential Imaging (ADI) also employed (see Section 4.3), allowing for individual  $Q$  and  $U$  images to be combined across a sequence (Column D). The square root of the sum of the squared  $Q$  and  $U$  images, is called the “Polarized Intensity” (PI) image (Column E). As can be seen in the figure, it easily isolates the (polarized) light of the disk from the (unpolarized) starlight. The combined total intensity image, on the other hand (Column E, bottom), is dominated by starlight.

different indices of refraction for different linear polarization states. If a single Wollaston is used, the light is split into two beams with orthogonal polarizations (often called the “ordinary” and “extraordinary” beams), while a double Wollaston will yield four beams, adding redundancy that helps in removal of detector location-specific artifacts. The precise orientation of the orthogonal ordinary and extraordinary polarization vectors relative to the sky is manipulated to fully sample the polarized emission from the source by rotating an optic called a half- or quarter-wave plate, which modulates the orientation of the linear polarization state of incoming light for the two channels. This modulation (generally sequences of 4 angles—0°, 22°5, 45°, 67°5) allows the images to be combined to yield the

Stokes polarization vectors  $I$ ,  $Q$ , and  $U$ .<sup>9</sup> Addition of images with orthogonal polarization captures the *unpolarized* intensity of the star, while subtraction yields either “ $Q$ ” or “ $U$ ” images, depending on the orientation of the wave plate.  $Q$  and  $U$  images are combined to isolate polarized light from the source via the equation  $PI = \sqrt{Q^2 + U^2}$ . Each sequence of wave plate angles thus produces four images—I,  $Q$ ,  $U$ , and PI.

<sup>9</sup> The Stokes vectors (a/k/a “Stokes parameters”) are a mathematical formalism used to describe the polarization state of light, namely: its total intensity ( $I$ ), its linear polarization state ( $Q$  and  $U$ ), and its circular polarization state ( $V$ ). HCI instruments are not generally sensitive to the fourth Stokes vector  $V$ , so I will not discuss it here.



The angle of the polarization vector can also be extracted from these quantities via the equation

$$\theta_p = 0.5 \arctan\left(\frac{U}{Q}\right)$$

The  $\theta_p$  vectors, when overplotted on images of a scattered light disk, demonstrate a characteristic centrosymmetric pattern. This is because of the preferred geometry of the scattering process; the most efficient scattering occurs when a photon's electric field orientation ( $\theta_p$ ) is orthogonal to both the line of sight and the vector connecting the scattering dust grain and star.

Extraction of polarized signals is complicated somewhat by multiple scattering and the internal optics of the instrument. Internal reflections result in depolarization effects that vary with wavelength, incident angle, and the thickness and index of refraction of the optical components. This induces so-called “instrumental polarization,” which is typically estimated from observations of both unpolarized, disk-free stars and polarization standard stars. The simple picture of polarization presented above also assumes that each photon received was scattered by only a single small dust grain in the disk on its journey from star to disk to Earth. This is a reasonable assumption in many cases, but multiple scattering does occur, resulting in deviations in the centrosymmetry of polarization vectors and, in the characteristic pattern of positive and negative signal in  $Q$  and  $U$  images (often called a “butterfly” pattern because the symmetric positive/negative lobes look a bit like butterfly wings). The inclination of the disk (i.e., whether emission is “forward” or “back” scattered) also impacts the efficiency of scattering, as do grain properties such as size, composition, and porosity.

The most common variation on the process described above is to compute the so-called “azimuthal” or “local” Stokes  $Q$  and  $U$  vectors, often denoted  $Q_\phi$  and  $U_\phi$  (e.g., Monnier et al. 2019; de Boer et al. 2020) and defined as:

$$Q_\phi = -Q \cos(2\phi) - U \sin(2\phi)$$

$$U_\phi = +Q \sin(2\phi) - U \cos(2\phi)$$

where  $\phi$  is the azimuthal angle. This formulation has the advantage of concentrating signal with the expected polarization vector orientation into the  $Q_\phi$  image, while the  $U_\phi$  image becomes an estimate of the noise induced by multiple scattering and instrumental polarization.

## 4.2. Reference Differential Imaging (RDI)

The Reference Differential Imaging (RDI) technique utilizes images of stars other than the science target to subtract starlight from a target image and is shown schematically in Figure 7. It is an ideal approach when either (a) the PSF of a system is exceptionally stable, often the case for space-based observatories such as HST, or (b) the source being targeted has

extended, symmetric features (e.g., a circumstellar disk) that might be subtracted by more aggressive algorithms that rely only on images of the target star for reference (see next several sections). Reference PSF libraries for RDI generally consist of images of many other stars taken at the target wavelength and in the same observing mode (e.g., same coronagraph) with the same instrument. In the case where a large library of reference images is available (e.g., a large HCI campaign, a well-established space telescope instrument), just a subset of the most highly correlated images may be chosen to construct a PSF.

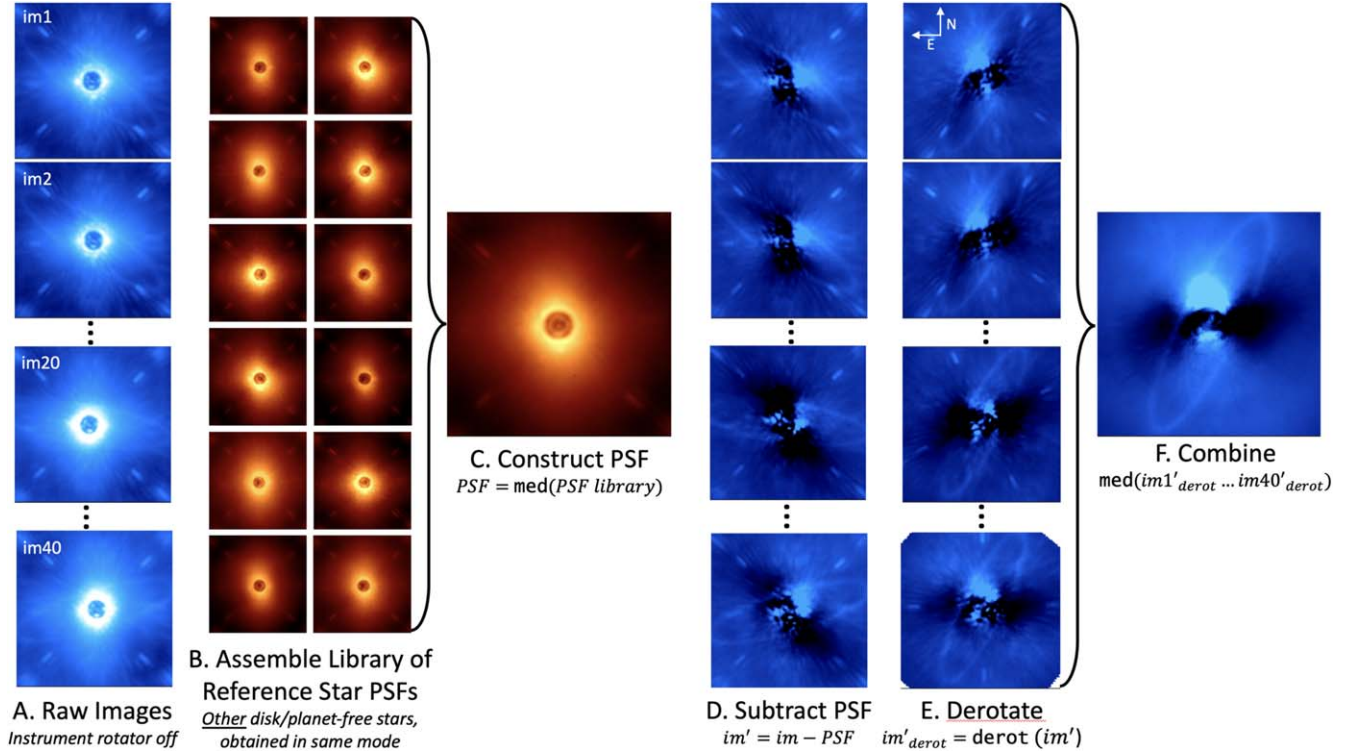
Some HCI observers, particularly of disks, regularly conduct PSF reference star observations as part of their efforts to observe a science target. PSF references are often chosen to be similar in location on the sky (so they can be observed interspersed with or immediately before or after the science target, at similar airmass), of similar apparent brightness at the wavelength of the WFS (so that the AO system performs similarly<sup>10</sup>), and of similar color (so that the science image(s) have similar properties).<sup>11</sup> Some modern HCI systems (SPHERE, MagAO-X) are equipped with “star-hopping” modes that allow the AO loop to be paused on one target (e.g., the science target) and then re-closed once the telescope is pointed at another nearby target (e.g., the PSF reference star). This ensures maximal similarity in their PSFs.

## 4.3. Angular Differential Imaging (ADI)

The Angular Differential Imaging (ADI) technique builds on the legacy of “roll-subtraction” pioneered with the Hubble Space Telescope (HST, e.g., Schneider et al. 2014). It leverages angular diversity to separate stable and quasi-stable PSF artifacts from true on-sky emission. ADI is predicated on the assumption that the instrumental PSF remains (relatively) stable in the frame of reference of the instrument throughout the image sequence, while true on-sky signal rotates with the sky. This allows the time series of images to be leveraged for pattern matching or statistical combination to estimate the stellar PSF and remove it. In practical terms, the quality of any ADI-based subtraction is a strong function of the amount of on-sky rotation of the source. For this reason, most direct imaging target observations are roughly centered around the time of that object's transit across the meridian, as this maximizes the amount of rotation achieved for a given amount of observing time. Rotation is essential to reduce a phenomenon called “self-

<sup>10</sup> One clever trick that some AO observers use is to “pause” the AO control loop, slew the telescope to the reference star, and re-close it with all the same WFS algorithmic parameters in order to maximize this similarity.

<sup>11</sup> Similarity in color is important in RDI primarily because WFS and detector wavelength ranges are often different. Ideally, the reference star should be of similar (or slightly higher, Debes et al. 2019) brightness at both wavelengths so that its total flux on the detector (at the science wavelength) and the performance of the AO system (set by the star's brightness at the WFS wavelength) are similar.



**Figure 7.** A schematic representation of the process of Reference Differential Imaging (RDI), in this case using Gemini Planet Imager *H*-band images of the debris-disk host HR4796A collapsed across all  $\sim 40$  wavelength channels of GPI. RDI utilizes a library of images of stars *other than the science target* (Column B) obtained in the same observing mode. Generally, stars without any known disk or planet signal are chosen as references. These reference images can be combined simply (e.g., median combined, Column C) or used to build a custom PSF for each target image in the sequence (see Section 5). This PSF estimate is subtracted (Column D) to remove starlight in the image. In the case where the images were obtained with the instrument rotator off (typical for ground-based observing, see Section 4.3), these subtracted images are rotated to a common on-sky orientation (Column E) and combined (Column F).

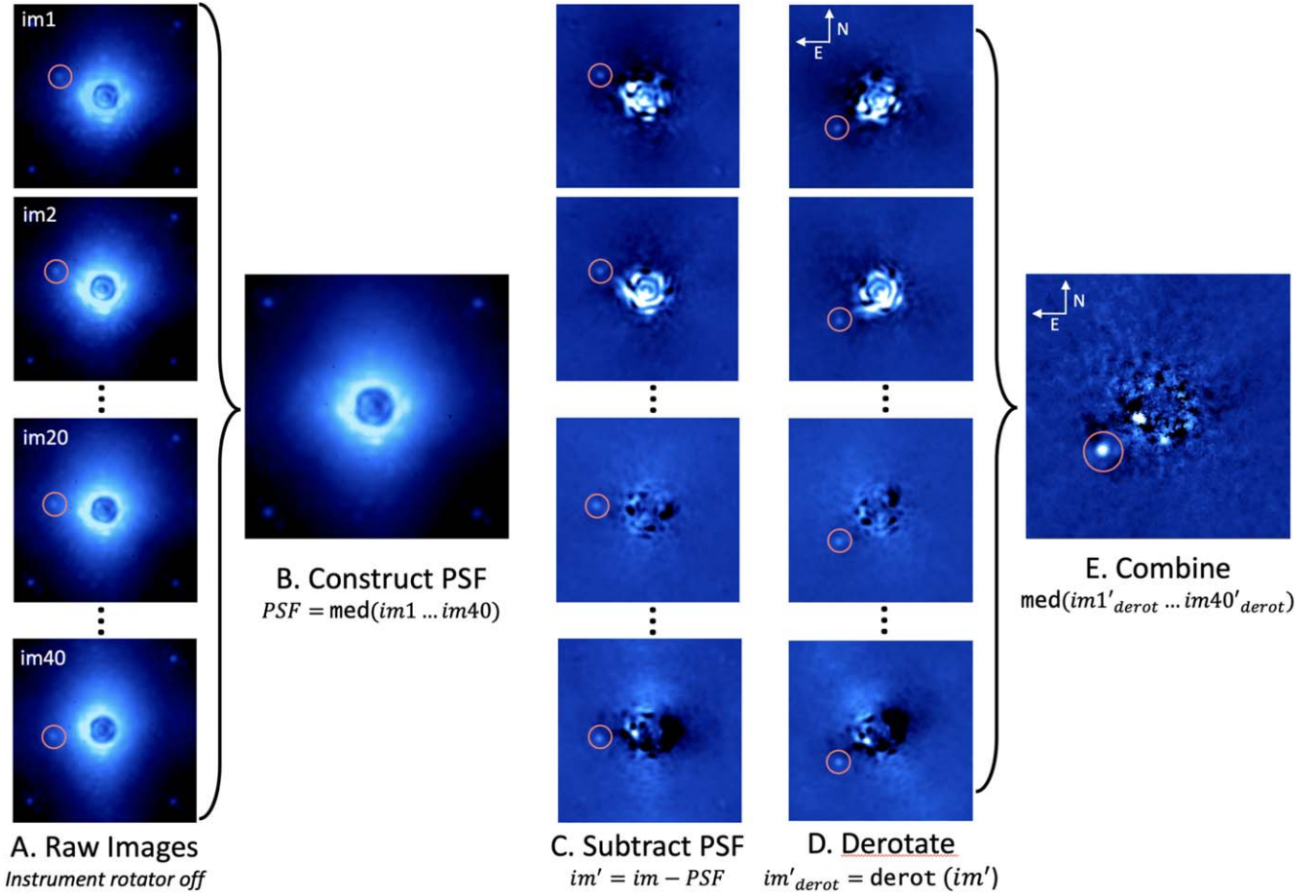
subtraction,” in which the signal of a source (disk or planet) is present in a different but nearby location in the PSF image being subtracted, resulting in characteristic negative lobes on either side of the source where it has been subtracted from itself (hence the name).

The simplest form of ADI, so-called “classical” ADI (cADI), constructs a single PSF for subtraction from the median combination of all images in a time series, subtracts this median PSF from each image and then rotates these subtracted images to a common on-sky orientation. These PSF subtracted and re-oriented images are then combined, further suppressing the residual speckle field, which varies from image to image.

#### 4.4. Spectral Differential Imaging (SDI)

HCI observing programs often aim not just to *detect* exoplanets and circumstellar disks, but also to *characterize* them, for which multiwavelength information is invaluable. Due to the many challenges of absolute photometric calibration in HCI (see Section 7), characterization is best facilitated by obtaining simultaneous imagery at multiple

wavelengths. Thus, many modern HCI instruments are so-called “Integral Field Spectrographs” (IFSes). IFS instruments are used throughout Astronomy with a range of architectures, but in the case of HCI, they are generally of a similar lenslet-based design. In lenslet-based IFSes, a grid of lenslets is placed in the focal/image plane of the optical system (not unlike the grid of lenslets placed in the pupil plane of a SHWFS, see Section 2.2), and the lenslet spots are dispersed to produce a spectrum for each lenslet. Each “spectral pixel,” or “spaxel” (also referred to as a “micro-spectrum”), contains spectral information *at a particular location in the image plane*. Microspectra are wavelength calibrated using observations of internal arc lamps, generally taken close in time to the science observations because the wavelength solutions are strongly dependent on instrument flexure. Raw IFS images are converted to multiwavelength image cubes by extracting photometry from the microspectra at specific wavelengths (using knowledge of the instrumental PSF). Photometric values for each extracted wavelength of the microspectrum are assigned to appropriate spatial



**Figure 8.** Illustration of the classical Angular Differential Imaging (cADI) technique. Images are derived from a sequence of 40 Gemini Planet Imager coronagraphic  $H$ -band ( $1.6\ \mu\text{m}$ ) images of the planet host Beta Pictoris ( $t_{\text{exp}} = 1$  minutes). Images (Column A) are collected with the instrument rotator off. The instrumental PSF (including any quasi-static speckles) remains relatively stable in the instrument frame throughout the sequence, while real sources rotate with the sky. The image sequence is median combined to create an instrumental PSF (Column B), which is then subtracted from each image (Column C), derotated to a common on sky orientation (Column D), and median combined again (Column E). In this case, the planet Beta Pictoris b (coral circle) is bright enough to be seen in individual exposures. The median PSF is not a perfect PSF reference, and image-to-image variation can be seen in Column C. However, derotating and median combining these imperfect subtracted images results in a very clear detection of the planet.

locations relative to those from neighboring microspectra in synthetic images. A raw IFS HCI of the planet-host Beta Pictoris is shown in Figure 9.

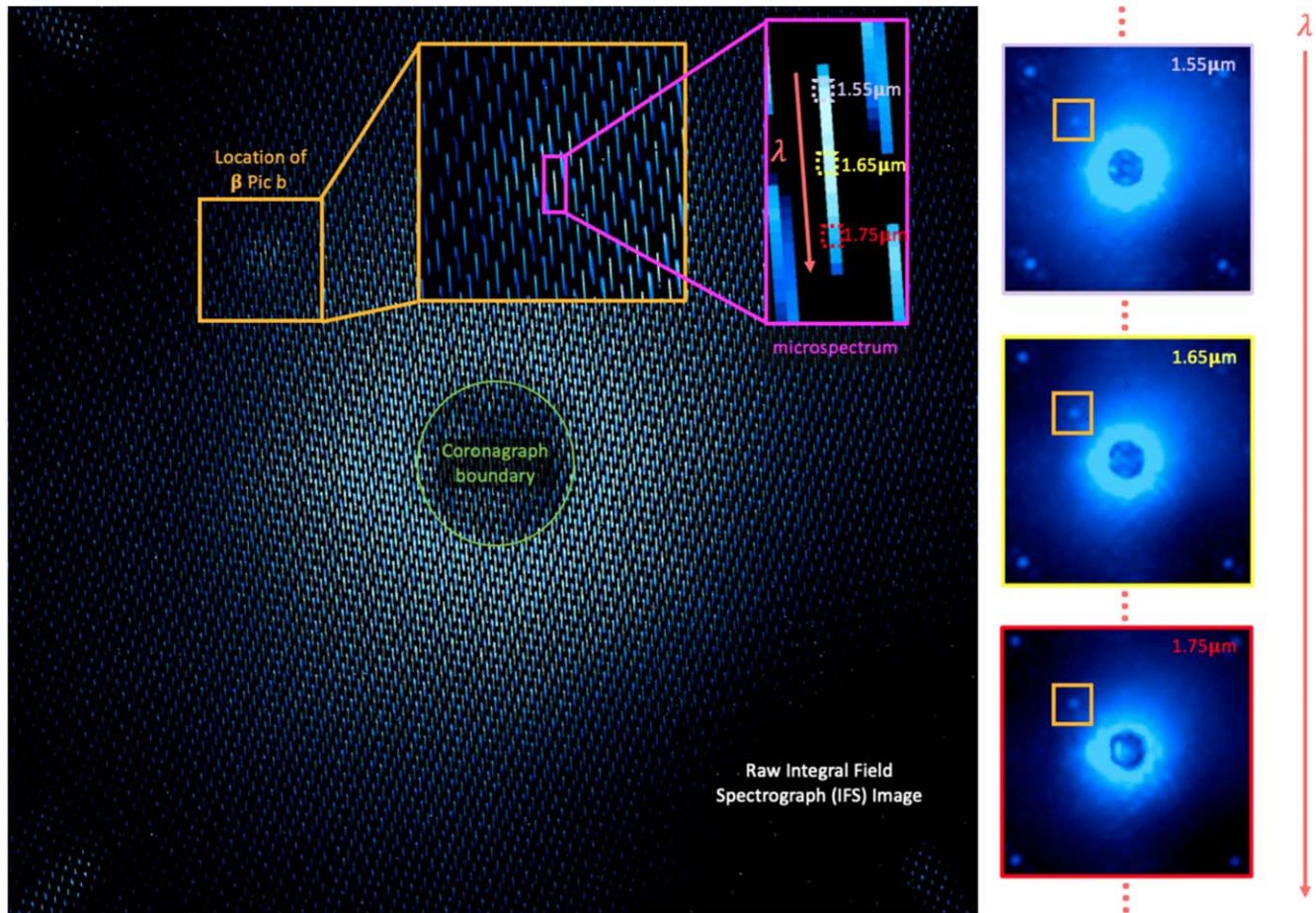
Spectral Differential Imaging (SDI) takes advantage of differences in the spectral properties of planet and star. In particular, it leverages images at wavelengths where planets are dim (e.g., for methane dominated planetary atmospheres, at  $1.5$  and  $1.7\ \mu\text{m}$ ) to construct a PSF model that is largely uncontaminated by planet light, limiting self-subtraction. Because images are collected at multiple wavelengths contemporaneously, this circumvents some of the effects of a temporally varying PSF. As a result, the library of reference images is often better matched to the target PSF.

Most high-contrast SDI imaging to date has been done with an Integral Field Spectrograph such as GPI or SPHERE. SDI can

also be implemented without an IFS by simply splitting incoming light into two beams with a 50/50 beamsplitter, dichroic, or Wollaston prism,<sup>12</sup> and passing each beam through a different narrowband filter. This is sometimes called *Simultaneous Differential Imaging* (still SDI). The filter pairs lie on- and off- of a spectral line of interest, and the most common lines used in today's high-contrast imaging campaigns are on- and off-methane in the NIR and on- and off-  $H\alpha$  in the optical. In the case of young moving group stars (ages 10–300 Myr), it is expected that planets

<sup>12</sup> A 50/50 beamsplitter splits light equally across a wide wavelength range. A dichroic is transmissive for some wavelengths and reflective for others, resulting in preservation of all of the intensity at a given wavelength. A Wollaston prism is similar to a 50/50 beamsplitter for the case of unpolarized input light—it does not split light by wavelength, but rather by polarization state.





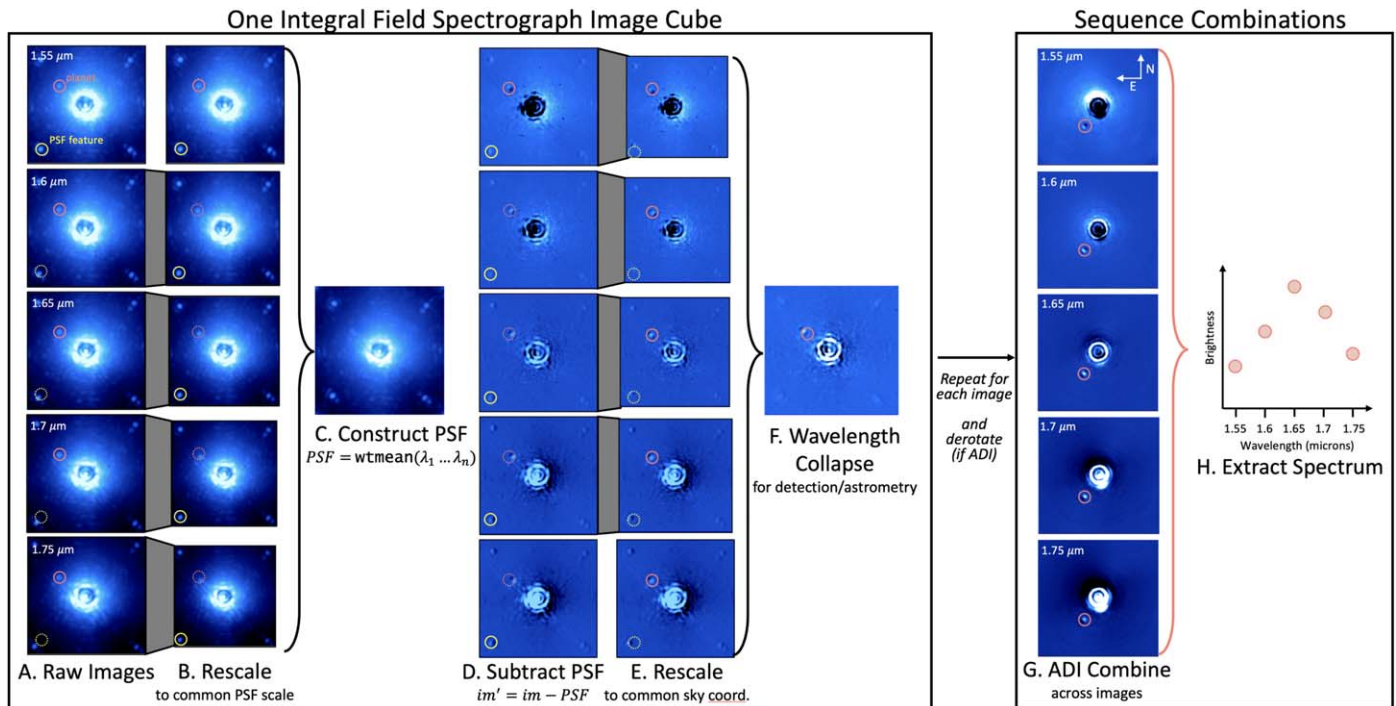
**Figure 9.** Schematic representation of the process of extracting a multiwavelength image cube from a single raw Integral Field Spectrograph (IFS) image. In this case, the background image is a raw  $H$ -band image of the star Beta Pictoris collected with the Gemini Planet Imager (GPI). Beta Pictoris has a known planetary companion, Beta Pictoris b, whose light can be seen even in raw GPI images as a region of excess brightness in the wings of the stellar PSF, indicated in orange here. IFS instruments place a grid of lenslets in the focal plane, and light from each is passed through a dispersing element before reaching the detector. This creates an array of microspectra on the detector, one of which is highlighted in magenta here. Each microspectrum can be wavelength calibrated using arc lamps and its brightness extracted to create a single spectral pixel, or “spaxel” for each wavelength (representative wavelengths of 1.55, 1.65, and 1.75  $\mu\text{m}$  indicated in cyan, yellow, and red on the microspectrum) and location in the image plane. These spaxels can be stitched together algorithmically to produce simultaneous images of the star at a number of wavelengths, creating a multiwavelength image cube rather than a single broadband image.

with methane-dominated atmospheres will be faint or undetectable in the methane band and brighter outside of it (see Figure 10).  $H\alpha$  differential imaging, on the other hand, leverages the fact that many younger ( $<10$  Myr) systems show evidence of ongoing accretion onto their central stars. The accreting material originates from and is processed through the circumstellar disk, meaning that any planets embedded in that disk are also likely to be actively accreting. One principal escape route for the energy of infalling material is radiation in hydrogen emission lines, particularly  $H\alpha$ , and we expect accreting protoplanets to be bright at this wavelength and faint or undetectable in the nearby continuum.

In terms of its utility as a tool to separate star and planet light, in its most generic form (what we might term “classical” SDI imaging, shown schematically in Figure 10) simply

leverages the fact that the physical size of a stellar PSF on a detector is a function of wavelength. For simultaneously acquired imagery at multiple wavelengths (i.e., a 3D cube of images with 2 spatial coordinates and 1 wavelength coordinate), this manifests as a magnifying effect as wavelength increases, and means that PSF features shift radially outward in detector coordinates, while true on-sky objects remain at the same position regardless of wavelength. Much like ADI angular rotation, the size of this effect is well-known (having a  $\lambda/D$  dependence), therefore it can be compensated for in post-processing. By expanding shorter wavelength images or compressing longer wavelength ones so that all simultaneously-obtained images share a common PSF scale, wavelength-independent features of the PSF can be estimated. This





**Figure 10.** A schematic representation of the process of “classical” Spectral Differential Imaging (SDI). Simultaneous images of a star are obtained at a range of wavelengths, in this case IFS images of the star Beta Pictoris obtained with the Gemini Planet Imager at *H*-band (1.5–1.75  $\mu\text{m}$ ). A representative set of 5 of 37 total wavelengths from the 3D image cube (2 spatial, 1 wavelength dimension) are shown in Column A, spanning a majority of the wavelength range. Each image is rescaled to compensate for the magnification of the stellar PSF with wavelength (Column B), placing instrumental PSF features on the same spatial scale (e.g. satellite spots, one of which is indicated in yellow throughout). This rescaling, however, shifts the position of any real on-sky signal (such as the light from the planetary companion Beta Pic b, indicated in pink throughout). Rescaled images can be combined (Column C) to create a relatively planet-free PSF (in this case by taking the weighted mean of the first and last few images in the rescaled image cube, where the planet light is farthest apart) and subtracted from each rescaled image (Column D) to remove a majority of the stellar signal. Rescaling must then be reversed (Column E) to re-align true on-sky signals before combination. Images can be combined in wavelength space to achieve detections or astrometric measurements (Column F), or the separate wavelengths can be retained and combined across a sequence of IFS images (Column G). Photometry of the planet can then be extracted from combined images to construct a spectrum (Column H).

rescaling alters the position of real objects in the images so that they are no longer in precisely the same location at all wavelengths, thus the rescaled images can be combined (e.g., via median or weighted-mean combination) to construct a relatively<sup>13</sup> planet-free PSF reference. This reference can be subtracted from rescaled images and the rescaling reversed to restore true on-sky coordinates, effectively realigning planetary signals across wavelengths. Multiwavelength image cubes can be collapsed in wavelength space to provide a robust planetary detection, enabling astrometric characterization. More commonly, however, wavelengths are kept separate and combined across a sequence of multiple IFS images. This enables extraction of planet photometry at each wavelength to create a coarse spectrum, with a spectral resolution controlled by how many spectral channels can be extracted from the microspectra,

<sup>13</sup> I say relatively only because the difference in the position of a planet in wavelength-rescaled images is generally small compared to typical angular rotations for ADI processing, and more planet light is likely to survive into any estimated PSF.

generally a few dozen over a  $\lesssim 0.5 \mu\text{m}$  wavelength range, for resolutions on the order of  $\sim 25\text{--}100$ .

SDI processing is rarely used in isolation or executed in the simple “classical” sense described above. Instead, it almost invariably applies more sophisticated PSF estimation techniques to create custom PSFs for each image within a sequence and each wavelength within the image cube (i.e., using KLIP or another algorithm). Combination of SDI and ADI processing allows the user to leverage both angular and spectral diversity to identify reference images where sources have moved enough to prevent their surviving into any combination (either through angular rotation or image rescaling).

In addition to taking advantage of the physical rescaling of the instrumental PSF, SDI processing also often involves the application of one or more planetary spectral templates to expand the reference library. For example, for a planet with a methane-dominated atmosphere, such as the planet 51 Eridani b, there are certain *H*-band wavelengths where methane absorption makes planetary signal undetectable. Regardless of their separation from the wavelength being subtracted, these

“planet-free” wavelengths can be leveraged to construct the PSF model.

The size of a stellar PSF is a function of wavelength; it increases as the wavelength does. Raw SDI image cubes are therefore not initially good references for one another. Their spatial scales must first be adjusted to a common magnification in order to construct a PSF library. While this makes the instrumental PSFs of the multiwavelength images match, a side effect of rescaling is that the true on-sky spatial scale varies across the wavelength dimension of the reference images. This often creates *radial* self-subtraction of the planetary PSF when planet light at another (rescaled) wavelength makes it into the library of reference images.

A distinct advantage of SDI is the acquisition of spectral information, which allows for atmospheric characterization of directly imaged companions and composition analyses of circumstellar disks. Although the mechanics of the technique are somewhat different and outside of the scope of this tutorial (relying on the placement of optical fibers on and off of the known location of a directly imaged companion), it is worth noting that medium- and high-resolution spectroscopy is increasingly used to finely characterize the atmospheres of directly imaged companions.

## 5. Algorithms for High-Contrast Image Processing

In addition to applying hardware (see Section 2) to suppress starlight and differential imaging (see Section 4) to facilitate separation of star and planet signal, most modern HCI efforts require additional post-processing beyond the “classical” versions described in Section 4, and the most common techniques a side effect of this are described in this section.

### 5.1. Filtering

A common form of preprocessing for high contrast images is the application of so-called “high-” or “low-pass” filters to the data. This terminology refers to the spatial frequencies<sup>14</sup> that are least suppressed by the filtering algorithm—they “pass through” the process relatively unscathed, while other spatial frequencies are suppressed. A highpass filter “passes” high spatial frequency signals such as narrow disk features and planets. A low-pass filter suppresses these signals while preserving extended structures such as the stellar halo or broad disk features.

Highpass filters can be applied to high-contrast imaging data before or after PSF subtraction. A simple example of a highpass filter is the so-called “unsharp masking” technique,

<sup>14</sup> This is a Fourier analysis term, and can be understood through the relation between pupil and image plane discussed previously. When an image undergoes Fourier transform, the intensity of the resulting 2D function can be related to the strength of various “spatial frequencies” in the image. These can be thought of as maps of the degree of symmetry and typical size scale of variations in the intensity of the image.

wherein an image is convolved with a simple kernel (often a Gaussian), and then this smoothed image is subtracted from the original. High spatial frequency structures are drastically altered (spread across many more pixels than their original extent) by this convolution, while low spatial frequency structures remain largely unaltered. Thus, subtraction of the smoothed image suppresses these low-frequency signals while preserving high-frequency structure. There are a range of additional algorithms used to achieve highpass filtering, many of which are applied to the Fourier transform of an image in the frequency domain all designed to serve the same purpose.

### 5.2. PSF Post-Processing

A number of post-processing algorithms extend the concept of “classical” differential imaging to construct *custom* PSF models for *every* image in a time series *individually*, rather than adopting a single representative PSF for the entire image sequence. Two of the most commonly used algorithms are outlined below. Like ADI, RDI, and SDI, both of these algorithms rely on assembly of a library of reference images (often other images of the target itself taken in the same imaging sequence), and reference images are used to construct the PSF model(s) for the target image. The quality of PSF models relies on the strength of correlations between the target image and the other images in the reference library, with the algorithms weighting most heavily reference images that are most closely correlated with the target image.<sup>15</sup> In this way, these algorithms are able to capture the time varying nature of the PSF and quasi-static speckles rather than relying on a single PSF for the entire image sequence. PSFs can be constructed for an entire image, or for azimuthally and/or radially divided subsections of the image, and these algorithms can be applied for ADI, SDI, RDI, and occasionally even PDI image processing.

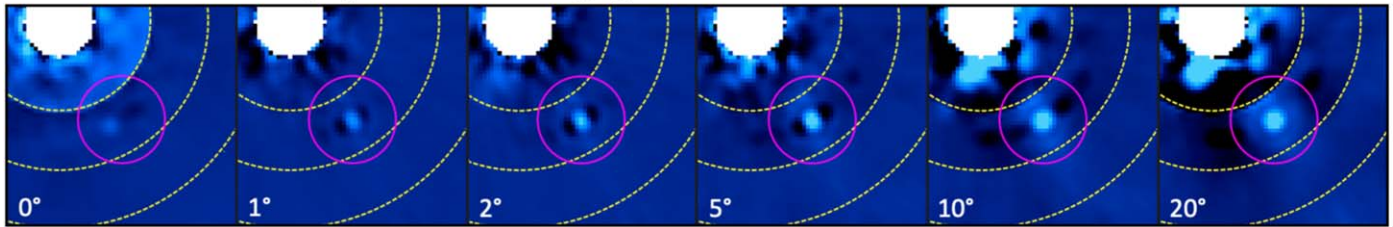
For these more advanced PSF-subtraction algorithms, restrictions are placed on which reference images are used to estimate the PSF for a given target image. The specific images in the sequence that are excluded and included in the reference library will change for each target image. Exclusion of images taken near in time or wavelength limits the amount of planet light that survives into the PSF model. The consequence of planet signal appearing in the PSF models is azimuthal (ADI) and/or radial (SDI) self-subtraction, as illustrated in Figure 11.

#### 5.2.1. KLIP

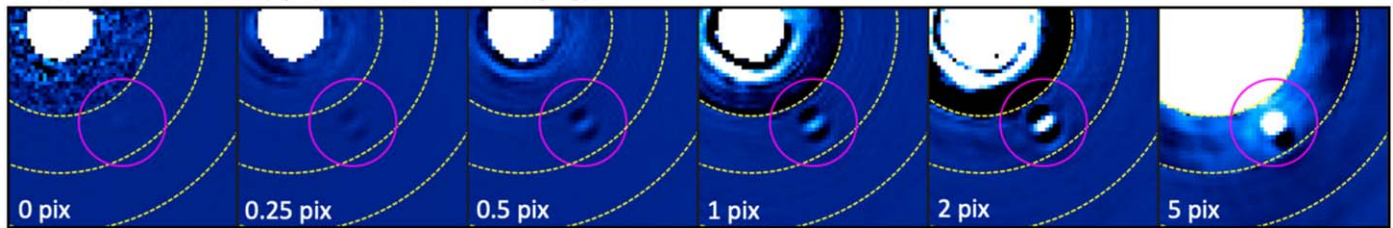
Karhunen Loeve Image Processing, or KLIP, is a statistical image processing technique in which images are converted to

<sup>15</sup> In the case of RDI processing of a disk-hosting star, some portion of the image known to host disk signal may be excluded from consideration (masked) before computing these correlations. This ensures that regions of relatively pure stellar signal drive the choice of reference images for PSF model construction and minimizes oversubtraction of disk signal.

## Azimuthal Self-Subtraction (Angular Differential Imaging)



## Radial Self-Subtraction (Spectral Differential Imaging)



**Figure 11.** A demonstration of varying degrees of azimuthal (top row) and radial (bottom row) self-subtraction of the planet Beta Pictoris b in KLIP-processed Gemini Planet Imager data. Azimuthal self-subtraction occurs in Angular Differential Imaging (ADI) when reference images where the planet’s signal fully or partially overlaps its location in the target image are included in the PSF reference library. Radial self-subtraction occurs in Spectral Differential Imaging (SDI) when rescaled (to match the scale of the target image) PSFs at nearby wavelengths contain planet signal (shifted inward or outward in the rescaling) that overlaps that of the target image. KLIP includes a threshold for the amount of angular or physical motion that a planet at a given location must undergo (due to angular rotation for ADI and PSF rescaling for SDI) before another image in the sequence can be included in the reference library for PSF subtraction. This is a tunable parameter, and both top and bottom panels depict a sequence of very aggressive (no threshold) to less aggressive reductions. An aggressive threshold generally provides better PSF subtraction (most evident at the center of the images) because the PSF library includes the images taken closest in time to the target image, but it also results in the highest degree of self-subtraction, evident in the overall suppression (faintness) of the post-processed planetary PSF and its narrowness compared to the much brighter and rounder planetary PSFs at right. The dark regions extending azimuthally (top row) or radially (bottom row) on either side of the core are referred to as “self-subtraction lobes”. These negative lobes reflect the presence of the planet in the KL modes, and they extend farther away from the planetary core in cases where there is less self-subtraction. Despite the resulting planetary signal suppression, removal of stellar signal is more effective under aggressive conditions, and fainter planets nearer the star may only be resolvable with more aggressive reductions.

1D vectors and cross correlated with all other images in a time sequence. This application of Principal Component Analysis (PCA) allows for identification of common patterns (“principal components”) in the image cube.

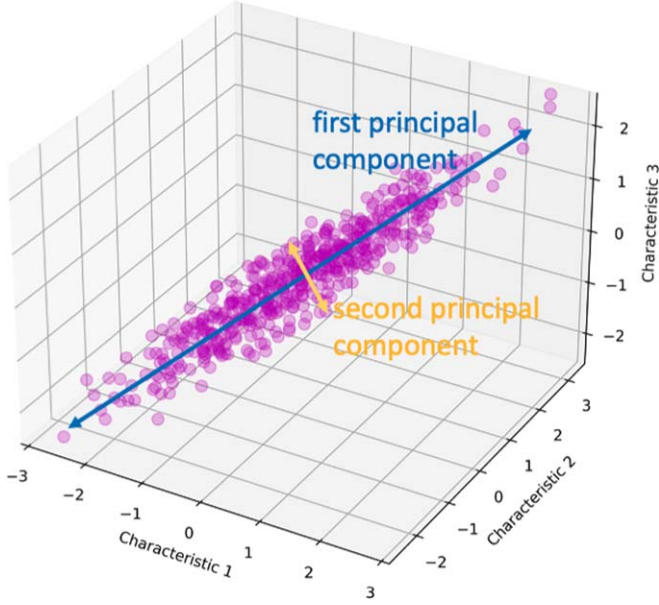
PCA is used in a range of contexts inside and outside astronomy to reduce the dimensionality of data. A simple example of how it works is to imagine a 3D scatterplot with evident correlations among the  $x$ ,  $y$ , and  $z$ -axis quantities (as shown in Figure 12). The  $x$ ,  $y$ , and  $z$  coordinates are, in such a case, not particularly good descriptors of the overall data, in that it is only in combination that they can describe its variation. If we were to instead define a first “principal component” axis along the line of best fit, this single variable would capture the most distinct first order pattern in the data (it is the best single descriptor of the data’s variance). If we were to add a second, perpendicular axis (in PCA each principal component is required to be orthogonal to all others), it would point in the direction of maximum scatter off the line of best fit, a good second order descriptor of the variance in the data.

It’s difficult to extend this toy example conceptually into high numbers of dimensions, but the principal is the same—each additional orthogonal vector must be orthogonal to all others and is chosen to describe the maximum amount of

additional variance in the data. Conceptually, in the case of PCA for HCI applications, this corresponds to patterns across many pixels that are present in the target image and some number of reference images. The first few principal components generally contain large scale PSF structures like core and halo, and the highest order principal components look like different realizations of the speckle pattern. Adding components to the model therefore increases its “aggressiveness.” This makes the likelihood of a well-matched PSF model higher, but also increases the likelihood that planet light will be oversubtracted or self-subtracted.

“KL modes” are the principal components of a library of reference images that have been transformed into 1D arrays (albeit with some complexities that I will not cover in detail here). Once they are computed, an individual image is “projected” onto these KL modes, which in practice looks like a weighted linear combination of the principal components. KLIP algorithms lend themselves easily to returning models of varying complexity (different numbers of KL modes) simultaneously, so PSF subtractions can readily be generated with a range of aggressiveness and compared. Low numbers of KL modes correspond to more conservative reductions, in that they (a) contain only the most widely varying PSF structures, and





**Figure 12.** A simple visualization of Principal Component Analysis (PCA). To describe the position of any one data point in this data set, one could specify three coordinates— $x$ ,  $y$ , and  $z$  location along the depicted axes. However, one could also provide a good approximation of a point’s location by specifying a single coordinate along a vector that describes as much of the variation in the data as possible—the so-called “first principal component” (depicted in blue here). If we also specified that point’s location along an additional vector defined to be both: (a) orthogonal to the first principal component, and (b) pointing along the (orthogonal) direction describing the greatest amount of additional variance in the data, this “second principal component” (depicted here in yellow), together with the first, would provide an even better estimate of the point’s location using only two (rather than the original three) coordinates. In high-contrast imaging, these patterns of covariance among images (principal components) can be used to model an image’s Point-Spread Function (PSF) using Karhunen–Loève Image Processing (KLIP), which is a variant of Principal Component Analysis.

(b) have relatively lower probability of any true circumstellar signals (disk, planet) being mistaken for PSF features. The probability of circumstellar sources being picked up in the KL modes is much higher for spatially extended disks than for planetary point sources, so KLIP-ed disk images often use a low number of KL modes (e.g.,  $<10$ ), while point-source reductions use dozens to hundreds. A schematic illustration of the KLIP process is shown in Figure 13.

### 5.2.2. LOCI

The Locally Optimized Combinations of Images (LOCI, Lafrenière et al. 2007) technique constructs a PSF model by weighting and combining some number of images from the reference library as a PSF model for the target image. In its original form, the algorithm computes a least-squares fit to the target image using weighted linear combinations of the images in the reference library, with the goal of minimizing the

residuals in the difference of the target image and the PSF model. Since it was originally developed, several enhancements have been made to the LOCI algorithm. A non-exhaustive list of these enhancements is provided below.

*Template LOCI*—(TLOCI, Marois et al. 2014), was specifically designed for SDI imaging and its aim is to maximize the SNR of planets with a given spectral shape. The user specifies a planet spectrum (e.g., flat, methane-dominated, etc.) and sets a threshold for the amount by which the planet’s flux is allowed to be reduced by self-subtraction (due to both azimuthal FOV rotation with time and radial PSF magnification with wavelength). Using simulated planets, the amount of self-subtraction in each reference image is quantified. Images with predicted self-subtraction above a certain threshold are excluded from the reference library before the least-squares fit is computed.

*Adaptive LOCI*—(ALOCI, Currie et al. 2012) implements an additional step of subtracting the radial profile of the star (the seeing halo) so that the speckle patterns among images can be readily compared. It also constructs a reference library from only the most correlated reference images (those above a certain user-defined correlation threshold).

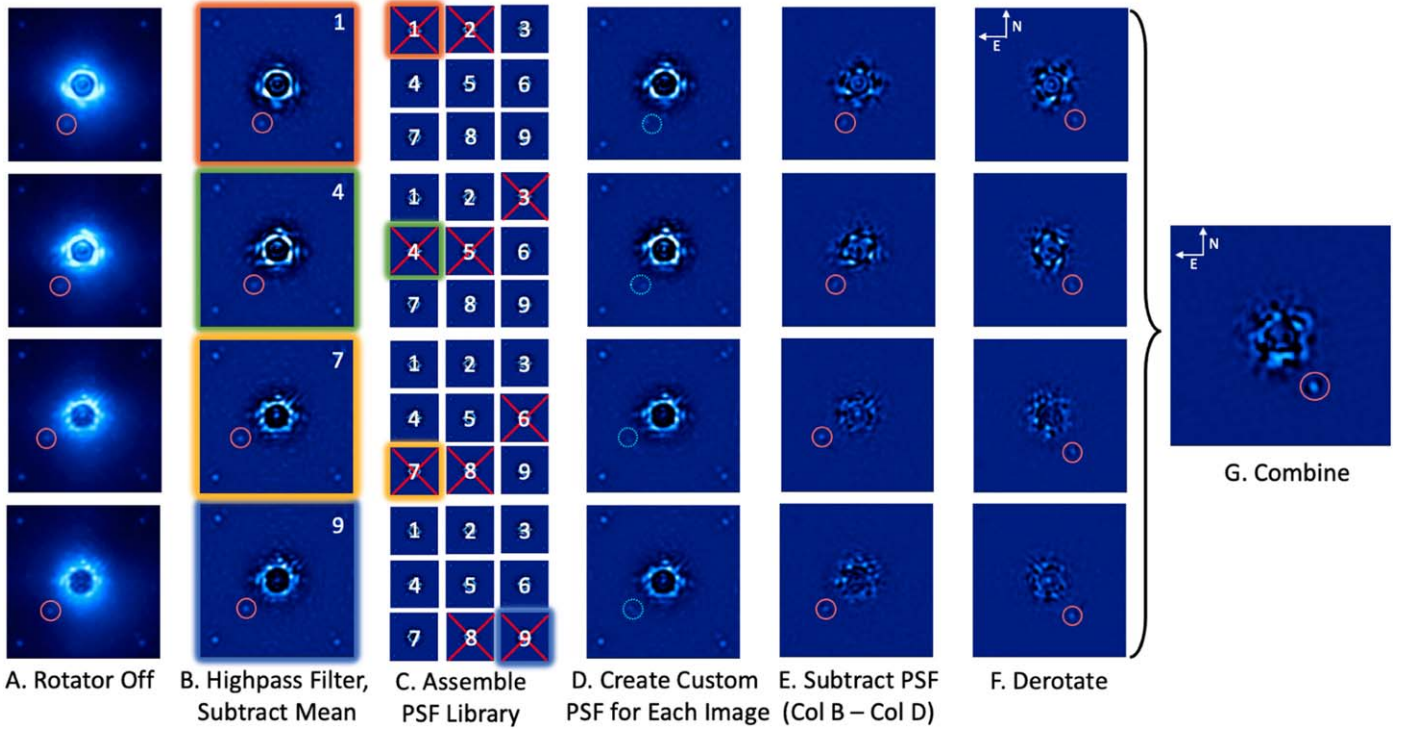
*The Signal to Noise Analysis Pipeline*—(SNAP, Thompson & Marois 2021) directly optimizes the nonlinear signal-to-noise equation for a planet at a given location by dividing the vicinity of a planetary signal into an annular “optimization region” and a smaller semi-annular “subtraction region.” Forward-modeled planet photometry, a vector of coefficients for the linear combination, and an estimate of the noise derived from those coefficients are optimized to maximize signal-to-noise ratio.

## 6. Comparison of Techniques

Now that we have introduced both differential imaging techniques generally and some of the processing algorithms that we use to extend them and isolate light from extremely faint circumstellar signals, we can compare the relative efficacy of and situations best suited to application of each technique. These considerations are summarized in Table 2. Another useful tool for comparing and contrasting techniques is examination of post-processed images generated with each technique for the same dataset. This is provided in Figure 14 using both a very faint planetary signal (that of 51 Eridani b, Macintosh et al. 2015) and a debris disk whose narrowness facilitates recovery under all of the algorithms (HR 4796A, Arriaga et al. 2020).

Both Table 2 and Figure 14 highlight the fact that choosing a technique requires consideration of many factors, including the feasibility of the observations and the specific science aims. An important takeaway is that differential imaging techniques can be especially powerful in combination. For example, recovery of a disk signal in both PDI and RDI or ADI allows for





**Figure 13.** Illustration of the Karhounen–Loeve Image Processing (KLIP) technique. This technique can be applied to ADI, SDI, and RDI imagery, but is shown for the ADI case here. Like Figure 8, this visualization utilizes a sequence of 40 Gemini Planet Imager coronagraphic  $H$ -band ( $1.6\ \mu\text{m}$ ) images of the planet host Beta Pictoris. Images (Column A) are collected with the instrument rotator off, allowing the sky to rotate. The images first undergo highpass filtering and mean subtraction (Column B) to remove extended structures (disk, stellar halo) and ensure that the images fed to KLIP have means of zero, respectively. A collection of other images in the sequence (Column C) are assembled for PSF modeling of each target image (here depicted as images 1, 4, 7 and 9 in a 9 image sequence). Algorithmic controls determine the degree of “aggressiveness” in including or excluding reference images taken near in time to the target image, where planetary signal may overlap (excluded images shown with red x symbols in Column C). Principal Component Analysis of the reference library and target image allows for construction of one or more PSF models of tunable complexity (number of principal components in the model, Column D depicts  $N = 5$  components). As in cADI, these models are subtracted from the target image (Column E), derotated to a common on sky orientation (Column F), and combined (Column G) to reveal the planet.

computation of the polarization fraction ( $P = PI/I$ ), a sensitive probe of the disk’s grain properties. For planets, recovery of signal via multiple processing techniques lends credence to its nature as a *bona fide* planet. In other words, the various techniques neither compete with nor supersede one another—often all are needed to construct a full picture.

## 7. Analysis of High-Contrast Images

### 7.1. Contrast Measurement

When reporting a high-contrast imaging detection, contrast is an important metric; however, it is also important in quantifying instrument performance in the case of a non-detection. Modern high-contrast imaging campaigns have surveyed a large number of young nearby stars with relatively few detections of exoplanets (e.g.,  $9^{+5}_{-4}$  planets for  $5\text{--}13M_{\text{Jup}}$  planets at separations of  $10\text{--}100\text{ au}$ , Nielsen et al. 2019), though they have been more successful at detecting circum-stellar disks (detection rates of  $\sim 30\%\text{--}100\%$ , depending on selection criteria, Esposito et al. 2020). One of the main

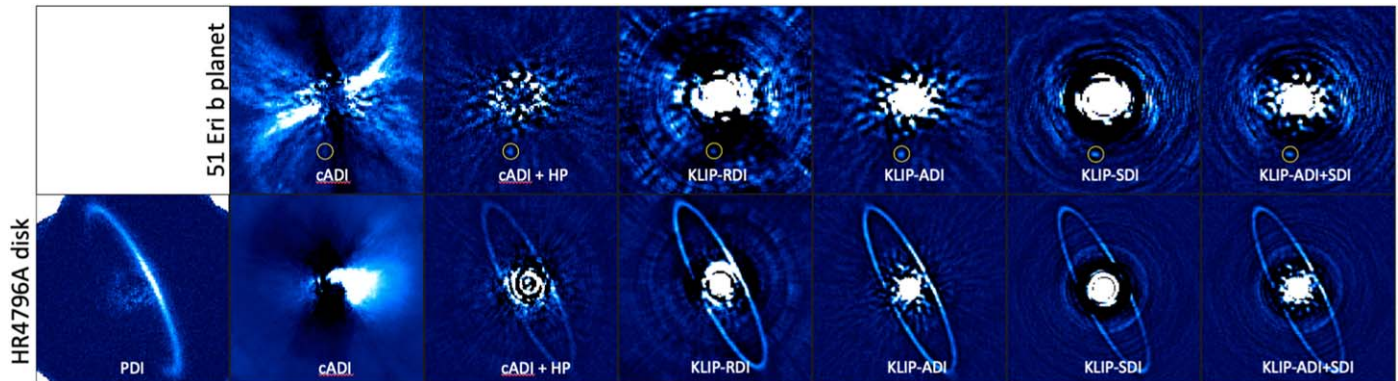
currencies of HCI surveys is therefore quantification of the instrumental performance, or limiting contrast, at a range of separations from each targeted star. This limiting contrast is a steep function of separation from the star, with lower contrasts achieved close to the star and higher contrasts at greater distances (see Figure 15). This means that a source at a given contrast is detectable in high-contrast images at a range of separations, with bright sources being detectable at all but the tightest separations and the faint sources only detectable far from the star.

“Contrast curves” therefore denote the detection threshold at each separation, with a few caveats and considerations. First and most importantly, many high-contrast imaging post-processing techniques (discussed in detail in Section 5) do not conserve the flux of astronomical sources. This means that the “raw” contrast, which is generally computed as 5 times the standard deviation of the noise at a given separation in the post-processed images, is not a true measure of the achieved sensitivity.

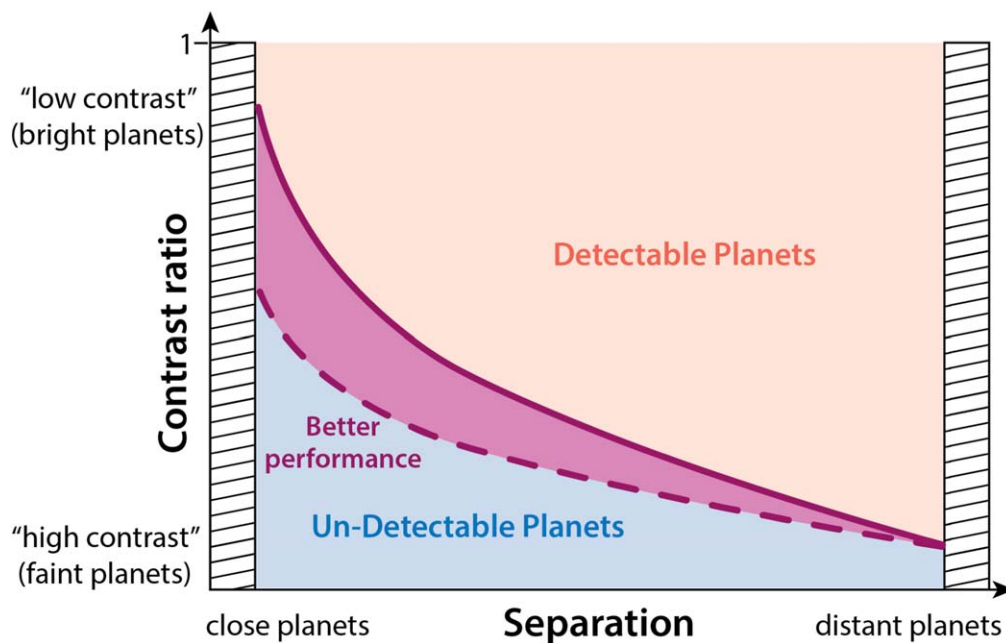
**Table 2**

Summary and Comparison of Differential Imaging Techniques, Highlighting the Variation in Requirements (Column 3), Science Aims (Column 4), and Pros and Cons (Columns 5 and 6, Respectively) Among Techniques

Technique	Abbr.	Requirements	Best for	Advantages	Disadvantages
Polarized Differential Imaging	PDI	Wollaston prism, rotating half wave plate	disk morphology and grain studies	<ul style="list-style-type: none"> <li>• does not require PSF subtraction</li> <li>• combined with total intensity imagery, probes disk grain properties</li> </ul>	<ul style="list-style-type: none"> <li>• instrumental polarization, multiple scattering effects difficult to isolate and remove</li> <li>• forward/back-scattering can result in only one side of a disk being detectable</li> </ul>
Reference Differential Imaging	RDI	reference star observations	detection and photometry of extended disks	<ul style="list-style-type: none"> <li>• allows for characterization of disks with arbitrary morphology, including face-on</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to achieve reference star observations with well-matched PSFs</li> <li>• PSF star observations require additional observing time</li> </ul>
Angular Differential Imaging	ADI	on-sky rotation	detection and photometry of planets, narrow disk structures	<ul style="list-style-type: none"> <li>• lots of on-sky rotation can enable more effective PSF subtraction close to star</li> </ul>	<ul style="list-style-type: none"> <li>• post-processed PSFs show azimuthal self-subtraction</li> </ul>
Spectral Differential Imaging	SDI	spectrograph	spectral characterization of planets, narrow disk structures	<ul style="list-style-type: none"> <li>• recovers spectral information, enabling characterization</li> <li>• can leverage knowledge/assumptions of spectrum to improve PSF subtraction</li> </ul>	<ul style="list-style-type: none"> <li>• post-processed PSFs show radial self-subtraction</li> <li>• planet movement constraint range is narrower</li> </ul>



**Figure 14.** A young exoplanet (51 Eri b, top row) and circumstellar disk (HR 4796 A, bottom row) reduced under a range of differential imaging techniques, from relatively conservative reductions at left to more aggressive reductions at right.



**Figure 15.** A schematic diagram illustrating how to read a contrast curve. At a given contrast and separation, a planet is detectable when it lies *above* the curve. Achieved contrast is a steep function of separation from the central star, with only the brightest planets detectable at tight separations.

In order to make a more accurate calculation, the algorithmic “throughput” must be computed by injecting sources into the image at a range of separations and quantifying their recovered brightnesses. Throughput is defined as the ratio of an object’s injected to recovered brightness (generally computed via the brightness of the peak pixel at the location of the source before and after PSF subtraction). Like contrast itself, it is a strong function of separation from the star. Throughput for most high-contrast imaging algorithms is low close to the star, meaning that source brightness is heavily suppressed by PSF subtraction, and approaches 1 at greater distances (meaning the

planetary signal is relatively unaltered by PSF subtraction). The best estimate of recoverable planet brightness is therefore the  $5\sigma$  noise level of the image divided by the instrument throughput at each separation from the star. This is sometimes called the “throughput-corrected” contrast, but is most often just referred to as “the contrast.”

When computing throughput, an important consideration is minimization of overlap/crosstalk between injected sources. As sources can overlap both azimuthally and radially, the general approach for quantifying detection limits has been to inject false planets in an outwardly spiraling pattern with

appropriate separations radially and azimuthally. This requires a choice of injected contrast for each false source. Generally, a low to moderate contrast is chosen and set uniformly throughout the injected planet spiral so that recovery is assured, however it is likely that injected object throughput is, at least to some extent, a function of brightness.

## 7.2. Limitations of Contrast as a Metric

Contrast curves have several limitations. First, they are sensitive to post-processing choices (e.g., KLIP parameters), therefore optimization can be computationally intensive. Second, they generally assume azimuthal symmetry in the sensitivity of post-processed images though, in reality, stellar PSFs often have azimuthally dependent structure. One common example of this is the so-called “wind butterfly” effect in AO-corrected images wherein lobes of higher noise/lower contrast are apparent on either side of a star in the direction of the wind. This means that neither noise nor algorithmic throughput is truly azimuthally symmetric. One way to mitigate this is to inject false planetary signals at various locations azimuthally and average their throughputs. For example, one might inject three spirals of false point sources with the spiral clocked by  $120^\circ$  each time in order to sample azimuthal variation in throughput.

A further complication is in the definition of the “noise” in post-processed images. The most typical noise estimate is the standard deviation of the residuals in the post-PSF subtracted image computed in small concentric annuli extending outward from the star. The convention in high-contrast imaging is to consider sources whose peak recovered brightness is at least 5 times above the noise level to be robust detections, and objects in the  $3\text{--}5\sigma$  range to be marginal. Many contrast curves reported in the literature are so-called “ $5\sigma$ ” contrast curves, but 3 or even  $1\sigma$  curves are also sometimes reported. One must be careful to understand and correct for any differences when comparing contrasts among surveys.

One final consideration in computing and interpreting noise in a post-processed image is that the dominant noise source close to the star is stellar speckles. In this speckle-dominated regime, there is a strong correlation between flux in adjacent pixels, since the stellar PSF has a width of several to many pixels. This has led to a best practice of implementing  $t$ -distribution rather than Gaussian noise statistics at tight separations, accounting for the small number of independent samples close to the star. In practice, this means dividing the computed standard deviation at a given separation by the factor  $\sqrt{1 + 1/n_2}$  (Mawet et al. 2014), where  $n_2$  is the number of independent noise realizations at that separation ( $\sim 2\pi r/\text{FWHM}$ ).

In summary, there are several important questions to ask oneself when studying a contrast curve.

1. Is it throughput corrected? If not, remember that the true limit is likely at lower contrast (a higher curve).
2. By what factor has the noise level been multiplied (1, 3, 5)? If less than five, recall that objects near the curve might be considered marginal or non-detections.
3. Has the noise level been corrected to reflect appropriate noise statistics near the star? If not, the true limit may be a steeper function of separation from the star than depicted.
4. How azimuthally symmetric is the post-processed image? If azimuthal structure is apparent, the curve should be interpreted as an average. In some parts of the image, objects below the curve may be detectable; in others, objects above the curve may be undetectable.

To put it plainly, the caveats described above mean that all contrast curves should be interpreted as relatively rough and fuzzy boundaries between detectable and undetectable planets. Furthermore, one must keep in mind when comparing contrast curves between studies and instruments that these choices may not be uniform among them and the curves may not be directly comparable. It is important when planning observations and interpreting detections (or non-detections) relative to contrast performance, to carefully read contrast curve descriptions and discern these important details. You may practice contrast curve comparison and parsing of these details by perusing Figure 16, which compares demonstrated and expected contrast for a range of current and future HCI instruments in several wavelength regimes.

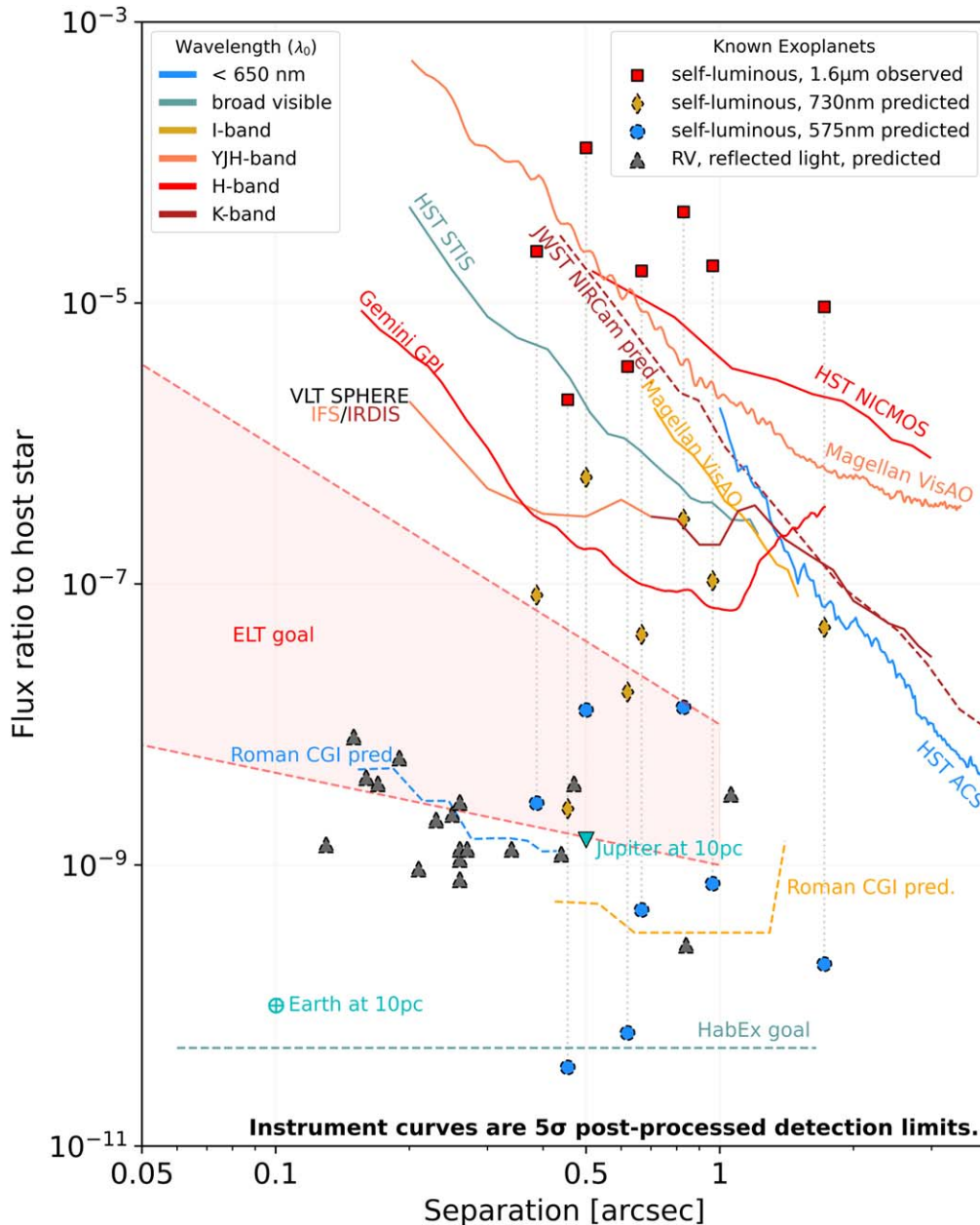
### 7.2.1. Aside: Contrast Curves for Disk Detections

Many of the points in the discussion above are altered or invalid for extended sources. Throughput, for example, is extremely difficult to compute for disks when their azimuthal and/or radial extent is large. Generally speaking, HCI disk detections utilize more conservative post-processing algorithms and observing techniques such as RDI for which throughput is much higher.

## 7.3. Signal-to-Noise Calculation

Signal-to-noise maps are standard in all fields of astronomy. In the case of direct imaging of point sources, there are several subtleties in computing them. First, the post-processed planetary PSF has characteristic “self-subtraction lobes” on either side of the planetary core. These are caused by the presence of the planet at different azimuthal angles in the reference library. The region containing the planetary core and self-subtraction lobes must be excluded in order to robustly estimate noise. This is typically done by masking this region and computing the standard deviation of the remaining pixels at a given radial separation. The nature of the speckle-dominated region of the PSF means that independent samples of the noise at a given radial separation are defined by the size of a speckle





**Figure 16.** Demonstrated (solid lines) and predicted (dashed lines) contrast performance of various current and future HCI instruments. Plot elements are color coded by wavelength of observation. Points indicate both detected (solid outline) and simulated (dashed outline) planets. (*code and data source*: V. Bailey) (Alphabetical descriptions of plot elements. DI: Selected self-luminous Directly Imaged (DI) exoplanets with known  $H$ -band contrasts. Predicted fluxes at Bands 1 and 3 are either from B. Lacy or COND (Teff  $\leq 1200$  K) or BT-SETTL models. Earth, Jupiter: simulated at quadrature as seen from 10 pc. (Jupiter albedo: 0.52 Traub & Oppenheimer 2010) ELT goal: Possible range of near-IR post-processed detection limits for next generation extremely large telescopes. GPI:  $5\sigma$  post-processed [KLIP + forward model match filter] contrast curve for  $H$ -band IFS mode, 1 hr integration. Calculated from an 11 minute  $H$ -band IFS observation of Sirius. (source: B. Macintosh). HabEx: Goal  $5\sigma$  post-processed contrast. IWA  $\sim 2.5 \lambda/D$  @ 450nm; OWA  $\sim 32 \lambda/D$  @  $1 \mu\text{m}$  (source: B. Mennesson) HST ACS:  $5\sigma$  post-processed [simple image difference] contrast curve of  $2 \times 100$  s Arcturus observation in F606W with  $1''.8$  occulter. (source: J. Krist) HST NICMOS: Best  $5\sigma$  post-processed [KLIP + match filter] contrast curve for F160W-band from the HST ALICE program (source: Choquet et al. 2014) HST STIS: Bar5 coronagraph  $5\sigma$  post-processed [KLIP] contrast curve; 162 s exposure, bandpass  $\sim 200\text{--}1030$  nm. (source: STIS handbook) JWST NIRCAM: simulated  $5\sigma$  post-processed [roll-subtraction] contrast curve for F210M-band. The model observation consists of  $2 \times 1$  hr rolls ( $\pm 5^\circ$ ), with a 10 mas pointing uncertainty and a 10 nm differential WFE. On sky JWST performance indicates that the true limit is lower still (source: Beichman et al. 2010). Roman CGI narrow FOV: Modeled  $5\sigma$  post-processed [RDI, fpp = 2] contrast curve for Band 1 imaging of a  $V = 5$  G0V star with the HLC coronagraph. Integration time is 10,000 hr (source: B. Nemati) Roman CGI wide FOV: Modeled  $5\sigma$  post-processed [RDI, fpp = 2] contrast curve for Band 4 imaging of a  $V = 5$  G0V star with the SPC wide FOV coronagraph, based on OS9. Integration time is 10,000 hr (source: B. Nemati). RV: All planets from NASA exoplanet archive with a semimajor axis of  $0''.12\text{--}1''.4$ , mass  $> 0.25 M_{\text{Jup}}$ , and host star  $V$  mag  $< 7$ . Lambertian flux ratio assumes: radius =  $1 R_{\text{Jup}}$ , geometric albedo = 0.5, circular orbit, inclination =  $90.0^\circ$ , and angle of  $0^\circ$  from the ascending node. SPHERE:  $5\sigma$  post-processed [SDI] contrast curve for a  $\sim 1$  hr integration on Sirius. At separations  $< 0''.7$  the curve is for IFS YJH, while  $> 0''.7$  is IRDIS K12 (Source: Vigan et al. 2015).

(the PSF FWHM), leaving relatively few independent noise samples at tight radial separations and requiring  $t$ -distribution noise statistics (Mawet et al. 2014).

#### 7.4. Astrometric, Photometric, and Spectral Extraction

PSF-subtraction techniques, while powerful for isolating faint signals, complicate the extraction of accurate astrometry, photometry, and spectra from a detected object because the process of PSF subtraction does not conserve planet signal. A number of strategies are used to mitigate these complications and extract robust estimates of planetary photometric, astrometric, and spectral signals in HCI.

*False Planet Injection*—Injection and recovery of false planet signals in the image helps quantify the amount of planetary signal lost during image processing (as described in Section 7.1). This is used, in turn, to correct photometry and estimate the true intensity of planet light at a given wavelength. Similarly, false planets can be used to quantify astrometric and photometric uncertainties, often by injecting them into raw images and utilizing the statistics of their recovered versus injected locations and fluxes to quantify uncertainty on astrometry and photometry of the companion.

*Forward Modeling*—Injection of a model companion or disk into raw images and examination of its morphology, astrometry, and photometry in post-processed images, is known as “forward modeling.” The properties of these false planets (brightness, location, fwhm) or disks (extent, inclination, radial brightness distribution) are iterated upon and forward models compared to data. This process is essential in interpreting post-processed images, which suffer from both self-subtraction (see “Signal-to-Noise Calculation” above) and so-called “over-subtraction,” in which some of the planet or disk signal is flagged as noise and subtracted. Forward models are tuned by minimizing residuals in the difference of the PSF subtracted and forward modeled images. In many cases, models are injected not into the target image sequence, but into a reference image sequence or at a wavelength in the target sequence at which the target signal is absent or minimized. Post-processed signals are dependent on their azimuthal and radial location, and on the precise PSF, which is wavelength dependent, so neither of these techniques provides a perfect match. However, Pueyo (2016) showed that a post-processed PSF can also be modeled for a particular location mathematically, without altering the original images, by propagating a perturbation to the covariance matrix forward through the algorithm (KLIP or LOCI). This removes the problem of mismatch by constructing a forward model at the same location and wavelength, and the authors demonstrated its ability to boost the accuracy of spectral extraction. Inferences made via forward-modeling are, however, limited by our ability to accurately model the true planet or disk signal, which is particularly difficult for complex off-axis or time-varying PSFs and non-axisymmetric disk

structures. Nevertheless, post-processed PSFs, by virtue of our precise knowledge of their constructed photometry and astrometry, are powerful probes of the effects of PSF subtraction on the properties of real signals.

*Negative Planets*—Another technique for determining planetary flux and location is to inject *negative* false planets into the raw image sequence at the location of the planet candidate, effectively canceling its signal. Post-processed residuals are then minimized to determine a best fit. Although this results in robust photometry and astrometry estimates, arguably better than using forward modeling, it is computationally intensive, and uncertainties on this technique are harder to estimate. Often, observers assign error bars “by eye” to capture the range of values that result in good subtractions. For example, an appropriate flux scaling should result in near-zero residuals and not clear over- or under-subtractions (i.e., clear residual planetary excess or a clear residual negative signal at the planet location).

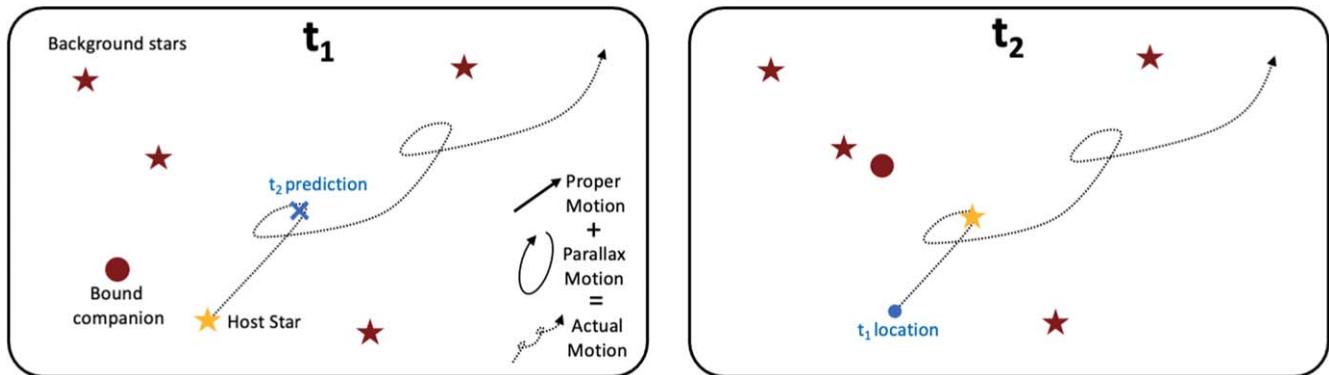
## 8. Potential Sources of False Positives

Direct imaging detections are intrinsically difficult, testing the limits of current technology, and there are a range of astrophysical and instrumental false positive possibilities.

### 8.1. Background Objects

One astrophysical false positive that can mimic a directly imaged companion signal is the coincidental alignment of a distant background source with a young star. This scenario is a possibility any time a faint point source is detected near a young star, thus it is among the first forms of vetting that all candidate planets are subjected to. For an initial single epoch detection, there are two important pieces of information used to assess the probability of a candidate being a background source—(1) the proximity of the target star to the galactic plane and (2) its spectrum. Coincidental alignments are much more common in the galactic plane, so the probability of false positives is higher in this case. As the most common background objects masquerading as planet candidates are distant red giants, spectral information—either true spectra or NIR colors—is also crucial in assessing the probability that a faint apparent companion is a young planet or brown dwarf.

With a few notable exceptions (e.g., 51 Eri, Macintosh et al. 2015, an object whose methane-dominated spectrum made its planetary nature clear from the outset), planet candidates are rarely announced until they have undergone an additional form of vetting—that of common proper motion with their host stars. Because the targets of direct imaging campaigns are close (generally  $<50$  pc), a necessity in order to achieve the requisite contrasts at planetary separations, their proper motions are invariably higher than those of distant background objects. Thus, most planet candidates are confirmed after obtaining a second epoch observation months or years after the initial



**Figure 17.** A schematic depiction of the process of determining common proper motion for a companion candidate (red circle) bound to a host star (yellow star). If the candidate is a true companion, then its motion over time (e.g., between epochs  $t_1$  at left and  $t_2$  at right) will closely follow the sky motion of the star (a combination of parallax and proper motion, shown as a dashed line). Companion-hosting stars are generally close to Earth, with a higher degree of proper motion and parallax than more distant background stars, which move very little between epochs. The orbit of the bound companion around the host star (not depicted here) can complicate this somewhat, but orbital motion is generally slow for the widely separated directly imaged companions detected to date. Importantly, color alone is rarely enough to determine whether a companion is bound or not, as background red giants share similar colors to directly imaged companions.

detection to confirm that the candidate and host star exhibit the same proper motions over that time period, as shown schematically in Figure 17. Candidates are ruled out as planets if they exhibit little to no proper motion between epochs.

In principle, establishment of common proper motion could be complicated by the additional motion of a true bound companion as it orbits its host star. In practice, however, most planet candidates are separated from their hosts by large enough physical separations that orbital motion is negligible compared to proper motion.

The most insidious form of false positive in establishing common proper motion is the coincidental alignment with the target star of an unbound foreground or background object with non-negligible proper motion. If the proper motion vectors of the two objects are in rough alignment and of similar magnitude, the time baseline needed to distinguish a comoving object is longer, as was the case with the apparent planetary companion HD 131399Ab (Nielsen et al. 2017).

## 8.2. Disk Features

Another form of astrophysical false positive results from the prevalence of circumstellar material around young stars targeted for direct imaging. Upon PSF subtraction, disk features can masquerade as planets, especially in cases where they are narrow (surviving highpass filtering) and non-axisymmetric. This is especially problematic for younger systems ( $<10$  Myr), where such features are ubiquitous (e.g., Benisty et al. 2023).

In the case of older ( $>10$  Myr) objects, for which the initial protoplanetary disk has usually either been incorporated into companions or dissipated, we see primarily second generation dust generated by the grinding of asteroids and/or comets in belts akin to our own asteroid and Kuiper belts. These belts

tend to be fairly symmetric and have limited spatial extent, making them much less likely to be confused for planet candidates. In known disk-bearing systems, candidate planets are vetted in several ways.

*Comparison with known disk features*—in both millimeter thermal emission and NIR scattered light (especially PDI-resolved features) informs the probability of confusion occurring at the location of a planet candidate. In cases where a candidate is well inside of a cleared cavity (e.g., PDS 70b, Keppler et al. 2018), the odds of confusion are minimal.

*Colors or spectra*—of companion candidate(s) can be compared to those of the star. In a case where the star and candidate spectra closely match, odds are good that the candidate is a scattered light feature. This could mean an envelope or disk around a planet, or a clump of disk material that has not yet formed a planet. In cases where a planet candidate exhibits a substantially different spectrum from that of the star, it is considered strong evidence for a planetary nature.

*Multiepoch information*—can distinguish static disk features from orbiting companions. Signals that show orbital motion are likely to be planetary, and those that are static are more likely disk features. This is complicated in the case of planet-induced spiral arms, which likely rotate with a pattern speed equal to the orbital speed of the companion inciting them. An important test is, therefore, whether apparent point sources that lie along spiral arms orbit with the speed of a companion at the point source’s orbital separation. If they orbit faster (or slower), this is consistent with incitement by a different planet on a closer (or more distant) orbit.

*The robustness of the signal among post-processing techniques*—particularly those that vary somewhat in “aggressiveness” is also an important part of vetting a planet candidate.



In the most insidious cases, the presence of an extended but narrow disk feature at different azimuths in the PSF reference library can lead it to appear point-like in post-processed images. Persistence of the feature across PSF subtraction algorithmic properties, and in particular its persistence across various HCI techniques, helps to distinguish this scenario from a true point-like source. In cases where the disk structures are well constrained (e.g., from PDI imaging), forward modeling can be used to understand the likely appearance of disk structures following PSF subtraction and compared against the images. RDI and cADI are considered the most conservative processing techniques, while LOCI-ADI and KLIP-ADI are more “aggressive” in that they tend to model smaller spatial scale PSF features, and model mismatch can therefore result in smaller spatial scale substructures mimicing planetary signals. Tunable parameters in the algorithms, such as the degree of rotational masking, the size of the regions for which PSFs are constructed separately, and the complexity/number of modes applied to construct the model, can also be altered to be more or less aggressive. For example, including images in the reference library that are close in rotational space (a small rotational mask), constructing custom PSFs for very small regions of images, and increasing the number of modes in the PSF model all represent more “aggressive” reductions that will effectively remove stellar signal, but will also increase the probability of false positives. These parameters should be iterated over to probe the robustness of apparent planetary signals.

Various optical artifacts, quasi-static speckles, cosmic rays, and speckle noise can also masquerade as planets in post-processed images. In general, the properties of such artifacts should not closely mimic those of true astrophysical sources (e.g., by demonstrating self-subtraction). Nevertheless, careful analysis of false alarm probabilities is important in conducting HCI, particularly for low SNR recoveries. The gold standard in candidate vetting remains multiepoch, multi-wavelength, multi-instrument observations of candidates demonstrating common proper motion with the host star and evidence of a non-stellar spectrum.

## 9. Other Related Technologies

Although this tutorial is focused specifically on ground-based, non-interferometric direct imaging techniques, there are several highly complementary techniques.

*Interferometric Techniques*—can be applied in HCI in several ways. First, the beams from multiple telescopes can be combined to both collect more light and achieve higher resolution than is achievable with a single telescope aperture (because the resolution of an interferometer is  $\lambda/2B$ , where  $B$  is the longest Baseline distance between telescopes). Even in the case where multiple telescopes are not available for use as a classical interferometer, a technique called “Non-Redundant Aperture Masking” (Nakajima et al. 1989) can be used to

achieve higher resolution with a single telescope. NRM requires the application of a pupil mask that is mostly opaque but contains a number of holes, each pair of which has a different separation and therefore probes a different spatial frequency. The maximum resolution achievable under this technique is half of the classical diffraction limit ( $\lambda/2D$ ), giving a distinct advantage at tight inner working angles for imaging companions. All interferometric imaging requires some degree of image reconstruction and is innately model-dependent, but these techniques open up additional discovery space at high spectral and/or spatial resolution.

*Space-Based HCI*—is another important complimentary technique. It shares many features with ground-based HCI, including the need for wave front sensing and control, image post-processing, and application of differential imaging techniques. Adaptive optics is unnecessary in space, though some space-based HCI concepts use much lower cadence active mirror control to correct for slower (e.g., thermal) drifts in the shape of incoming wavefronts. RDI is more powerful in space because of the stability of space-based instrumental PSFs, allowing in some cases for a reference library composed of images of tens to hundreds of sources. Although space-based telescopes cannot leverage the rotation of the Earth to accomplish Angular Differential Imaging, they can apply a similar technique called “Roll Subtraction” by rotating the telescope around its optical axis during an imaging sequence. The amount of achievable rotation and the number of reference angles in such cases is small (e.g., 2 reference angles separated by  $\sim 15$  deg), but has proven effective at enabling ADI in space. SDI and PDI techniques are more or less unchanged in the space-based imaging scenario, though there are no plans to include PDI capabilities on any near-future space-based HCI missions.

*Sub-mm Interferometry*—is unrelated to HCI, but is highly complimentary, particularly for understanding scattered light disk features and protoplanets. Interferometric sub-mm arrays (e.g. ALMA) probe thermal emission from large grains in the midplane of disks. Together with information from NIR HCI, which probes the surface layers of the disk, and millimeter emission from molecular gas species, a holistic picture of a disk system can be formed that encompasses all three key components—large grains, small grains, and gas. Very high-resolution millimeter continuum imaging can even probe the presence of circumplanetary dust and gas, compelling evidence for the presence of protoplanets.

## 10. Conclusion

Over the past fifteen years, ground-based High-Contrast Imaging has proven to be a robust and versatile way to probe the properties of young exoplanets and circumstellar disks. Using adaptive optics and wave front sensing/control algorithms, atmospheric scintillation can be sensed and corrected,

allowing large ground-based telescopes to achieve diffraction-limited or nearly diffraction-limited imaging at optical and near-infrared wavelengths. HCI instruments often to apply first-order suppression of incoming starlight with coronagraphy. Differential imaging techniques are then applied to leverage polarimetric, spectroscopic objects, and angular diversity to model and remove additional starlight. Finally, post-processing algorithms are applied to enable detection of faint planetary and circumstellar disk signals and detailed spectroscopic, photometric, and astrometric characterization of them. Planet candidates are vetted by demonstrating common proper motion with the host star, robustness to algorithmic parameters, consistency with forward models, diversity in polarimetric or spectral properties relative to their host stars, and/or persistence across epochs, wavelengths, and instruments. HCI instruments and reduction techniques are necessarily complex in order to overcome the tremendous contrast and angular resolution barriers required to directly isolate the light from exoplanets and circumstellar disks. Yet, these techniques provide the best future prospects for detecting and characterizing an exo-Earth.

This tutorial was designed as an introduction for beginners, and is not comprehensive in its technical details. My hope is that it will enable those just getting started in the field to access more technical HCI instrument manuals and published results. To learn more about the current state of the art in high-contrast imaging, please see [bit.ly/beginHCI](https://bit.ly/beginHCI), which provides a “Reading/Viewing List for Beginning High-Contrast Imagers.”

## Acknowledgments

I would like to thank the wonderful undergraduate and graduate students in my Spring, 2023 research group for the many group meeting sessions of figure critiques that they engaged in—this article is much better for their feedback. They are: Sarah Betti, Jada Louison, Cat Sarosi, Cailin Plunkett, Alyssa Cordero, and Adrian Friedman. Thank you to Kim Ward-Duong and Cat Sarosi for their thorough reviews of the text of the article, and to Bruce Macintosh, Mark Marley, Max Millar-Blanchaer, Ewan Douglas, Christian Marois, and Rob de Rosa for consulting on various parts of it. Thank you to the anonymous reviewer for their extremely constructive feedback, which greatly improved the article. Finally, a huge thank you to my team of “internal” student reviewers—Giselle Hoermann, Kinsey Cronin, Jessica Labossiere, and Jingyi Zhang. I also acknowledge the generous support of the National Science Foundation grant AST-2009816 and the Research Corporation for Scientific Advancement Cottrell Scholar Award #28290 for completing this work.

## ORCID iDs

Katherine B. Follette  <https://orcid.org/0000-0002-7821-0695>

## References

- Arriaga, P., Fitzgerald, M. P., Duchêne, G., et al. 2020, *AJ*, **160**, 79
- Avenhaus, H., Quanz, S. P., Garufi, A., et al. 2018, *ApJ*, **863**, 44
- Balmer, W. O., Follette, K. B., Close, L. M., et al. 2022, *AJ*, **164**, 29
- Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, *A&A*, **402**, 701
- Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, *A&A*, **577**, A42
- Batalha, N. E., Marley, M. S., Lewis, N. K., & Fortney, J. J. 2019, *ApJ*, **878**, 70
- Beichman, C. A., Krist, J., Trauger, J. T., et al. 2010, *PASP*, **122**, 162
- Benisty, M., Dominik, C., Follette, K., et al. 2023, in ASP Conf. Ser. 534, *Protostars and Planets VII*, ed. S. I. Inutsuka et al. (San Francisco, CA: ASP), 605
- Betti, S. K., Follette, K., Jorquera, S., et al. 2022, *AJ*, **163**, 145
- Biller, B. A., & Bonnefoy, M. 2018, in *Handbook of Exoplanets*, ed. H. J. Deeg & J. A. Belmonte (Berlin: Springer), 101
- Bordé, P. J., & Traub, W. A. 2006, *ApJ*, **638**, 488
- Bowler, B. P. 2016, *PASP*, **128**, 102001
- Brandt, G. M., Brandt, T. D., Dupuy, T. J., Li, Y., & Michalik, D. 2021, *AJ*, **161**, 179
- Chauvin, G., Lagrange, A. M., Dumas, C., et al. 2004, *A&A*, **425**, L29
- Chen, C., Mazoyer, J., Poteet, C. A., et al. 2020, *ApJ*, **898**, 55
- Currie, T., Fukagawa, M., Thalmann, C., Matsumura, S., & Plavchan, P. 2012, *ApJL*, **755**, L34
- Davies, M. B., Adams, F. C., Armitage, P., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 787
- de Boer, J., Langlois, M., van Holstein, R. G., et al. 2020, *A&A*, **633**, A63
- Debes, J. H., Ren, B., & Schneider, G. 2019, *JATIS*, **5**, 035003
- Esposito, T. M., Kalas, P., Fitzgerald, M. P., et al. 2020, *AJ*, **160**, 24
- Fehr, A., Hughes, A. M., Dawson, R. I., et al. 2022, *ApJ*, **939**, 56
- Guyon, O. 2018, *ARA&A*, **56**, 315
- Guyon, O., & Males, J. 2017, arXiv:1707.00570
- Kalas, P., Graham, J. R., Chiang, E., et al. 2008, *Sci*, **322**, 1345
- Keppler, M., Benisty, M., Müller, A., et al. 2018, *A&A*, **617**, A44
- Konopacky, Q. M., & Barman, T. S. 2019, in HR8799: Imaging a System of Exoplanets, ed. H. J. Deeg & J. A. Belmonte (Berlin: Springer), 1
- Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, *ApJ*, **660**, 770
- Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Sci*, **350**, 64
- Marley, M. S., Fortney, J., Seager, S., & Barman, T. 2007, in *Protostars and Planets V*, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press)
- Marley, M. S., Saumon, D., Visscher, C., et al. 2021, *ApJ*, **920**, 85
- Marois, C., Correia, C., Galicher, R., et al. 2014, *Proc. SPIE*, **9148**, 91480U
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Sci*, **322**, 1348
- Martinache, F., Guyon, O., Jovanovic, N., et al. 2014, *PASP*, **126**, 565
- Mawet, D., Milli, J., Wahhaj, Z., et al. 2014, *ApJ*, **792**, 97
- Mawet, D., Wizinowich, P., Dekany, R., et al. 2016, *Proc. SPIE*, **9909**, 99090D
- Monnier, J. D., Harries, T. J., Bae, J., et al. 2019, *ApJ*, **872**, 122
- Nakajima, T., Kulkarni, S. R., Gorham, P. W., et al. 1989, *AJ*, **97**, 1510
- Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, *AJ*, **158**, 13
- Nielsen, E. L., Rosa, R. J. D., Rameau, J., et al. 2017, *AJ*, **154**, 218
- National Academy of Sciences Engineering, N. A., & Medicine 2021, *Pathways to Discovery in Astronomy and Astrophysics for the 2020s* (Washington, DC: The National Academies Press)
- Poyneer, L. A., Macintosh, B. A., & Véran, J.-P. 2007, *JOSAA*, **24**, 2645
- Poyneer, L. A., Palmer, D. W., Macintosh, B., et al. 2016, *ApOpt*, **55**, 323
- Pueyo, L. 2016, *ApJ*, **824**, 117
- Choquet, É., Pueyo, L., Hagan, J. B., et al. 2014, *Proc. SPIE*, **9143**, 914357
- Rich, E. A., Monnier, J. D., Aarnio, A., et al. 2022, *AJ*, **164**, 109
- Roberts, L. J., Perrin, M., Marchis, F., et al. 2004, *Proc. SPIE*, **5490**, 504

- Schneider, G., Grady, C. A., Hines, D. C., et al. 2014, [AJ](#), **148**, 59
- Spiegel, D. S., & Burrows, A. 2012, [ApJ](#), **745**, 174
- Thompson, W., & Marois, C. 2021, [AJ](#), **161**, 236
- Traub, W. A., & Oppenheimer, B. R. 2010, in *Exoplanets*, ed. S. Seager (Tucson, AZ: Univ. Arizona Press), 111
- Vigan, A., Gry, C., Salter, G., et al. 2015, [MNRAS](#), **454**, 129
- Wagner, K., Apai, D., & Kratter, K. M. 2019, [ApJ](#), **877**, 46
- Wang, J. J., Rajan, A., Graham, J. R., et al. 2014, [Proc. SPIE](#), **9147**, 914755
- Winn, J. N., & Fabrycky, D. C. 2015, [ARA&A](#), **53**, 409