Kernel. Application and scientific potentials of full pupil interferometry : Statistical analysis of kernel observables

Alban Ceau

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES FONDAMENTALES ET APPLIQUÉES

THÈSE DE DOCTORAT

Kernel: Application et potentiels scientifiques de l'interférométrie pleine pupille

Alban Ceau

Laboratoire Lagrange, Observatoire de la Côte d'Azur

Présentée en vue de l'obtention du grade de docteur en Sciences de la Planète et de l'Univers **de l'**Université Côte d'Azur
Dirigée par : David Mary et Frantz Martinache

Soutenue le 30 juin 2020

Devant le jury, composé de :
Maud Langlois, Directrice de Recherches, Centre de Recherche Astrophysique de Lyon, Lyon 1

Frédéric Cassaing, Maitre de Recherches, Office National d'Études et de Recherches Aérospatiales

Karine Perraut, Astronome, Institut de Planétonolgie et d'Astrophysique de Grenoble

Romain Petrov, Directeur de Recherches, Laboratoire Lagrange, Observatoire de la Côte d'Azur

UNIVERSITÉ CÔTE D'AZUR

Abstract

Observatoire de la Côte d'Azur

Doctorat d'astrophysique

Kernel : Application and science potentials of full pupil interferometry

by Alban Ceau

Abstract (français)

Pour obtenir des observations à haute résolution angulaire précises (que ce soit en interférométrie ou en imagerie), il faut systématiquement passer par une quantité complexe qui correspond à la transformée de Fourier de la distribution spatiale d'intensité de l'objet. Pour toutes les techniques d'observation à haute résolution angulaire, la mesure de la phase de cette quantité complexe est limitée par la différence de marche introduites par l'instrument, notamment au sol, où la turbulence atmosphérique introduit des erreurs importantes. Ici, nous examinons des kernel phases, des observables peu sensibles à ces perturbations.

Dans ce cas, je me suis concentré sur la détection de binaires dans les kernel phases extraites à partir d'images, en utilisant des méthodes statistiques robustes. En théorie de la détection, la procédure la plus efficace pour détecter un signal dans des données bruitées est le rapport de vraisemblance. Ici, je propose trois tests, tous basés sur cette procédure optimale pour effectuer des détection systématiques de binaires dans des images. Ces procédures sont applicables aux kernels phases extraites à partir de n'importe quelle image.

Nous nous sommes concentrés sur des images de naines brunes froides du télescope spatial James Webb (JWST), afin de prédire les performances en détection de compagnons de ces objets. Jusqu'à maintenant, l'observation de ces naines, dites de type Y a été compliquée par leur faible température et luminosité, qui rend leur observation très difficile dans le proche infrarouge, le domaine de prédilection des observatoires au sol. Grâce à sa grande sensibilité et stabilité, JWST pourra observer ces objets avec une précision jusque là jamais atteinte. Grâce à cette stabilité, les images produites par JWST sont des candidates idéales pour des observations kernel.

Pour ces données, nous montrons que des détections de binaires sont possible à des contrastes pouvant atteindre 10^{-3} à des séparations correspondant à $1\lambda/D$, qui est communément admise comme la "limite de résolution“ d'un

télescope formant des images. Ces performances font de l'interférométrie kernel une méthode performante pour la détection de binaires de faible intensité. Ces limites dépendent fortement du flux disponible, qui détermine l'erreur sur les valeurs de flux mesurées au niveau de chaque pixel, et, par extension les erreurs qui affectent les kernel phases. Pour JWST NIRISS, dans le cas le plus favorable, c'est à dire la cible avec le plus haut flux pour laquelle il est possible de former des images sans saturation du détecteur, les contrastes auxquels il est possible de détecter un compagnon atteignent 10^{-4} à 1 λ/D avec un taux de faux positif inférieur à 1%.

Finalement, en annexe, je montre que la dégradation des limites de détection aux faibles séparation peuvent être attribuées à une dégénérescence entre la signature d'une binaire et une erreur de tip-tilt sur la phase en plan pupille, qui elle même correspond à un déplacement de l'image sur le plan focal. Comme cette signature est tuée par la transformation kernel, et correspond à un décalage de l'image, elle peut être considérée comme une limite fondamentale à la détection de compagnons à très faibles séparations ($> 0.25\lambda/D$).

Abstract (english)

High angular resolution observations (both in interferometry and imaging) are obtained by measuring a complex quantity, corresponding to the Fourier transform of the spatial intensity distribution of the observed object. For every high angular resolution observation technique, the measurement of the phase of this quantity is limited by the optical path difference introduced by the instrument, especially on the ground, where atmospheric turbulence leads to large errors. This thesis work is built around kernel phases, observables that are largely insensitive to these perturbations.

In this thesis work, I focused on the detection of binaries from kernel phases extracted in images, using robust statistical methods. In detection theory, the most efficient procedure to detect a signal in noisy data is the likelihood ratio. Here, I proposed three tests, all based on this optimal procedure to perform systematic detections of binaries in images. These procedures can be used on kernel phases extracted from any image.

The use case we chose to evaluate these techniques was images of cold brown dwarfs produced by he James Webb Space Telescope (JWST), to predict the detection performances of companions around them. Currently, observation of these cold, Y type dwarfs has been made difficult by their very weak luminosity and temperature, which make observing them very difficult in the near infrared, the preferred domain of AO corrected ground based observatories. Thanks to its great sensitivity and stability, JWST will be able to observe these objects with the greatest precision achieved yet. This stability makes images produced by this telescope ideal candidates for kernel analysis.

For these images, we show that binary detections are possible a contrasts that can reach 10^{-3} at separations corresponding to $1\lambda/D$, often considered to be the resolution limit of a telescope. These contrast detection limits make kernel interferometry a powerful method for the detection of low flux binaries. These detection limits strongly depend on the available flux, which determines the error level on each pixel, and therefore the noise that affects the kernel phases. For JWST NIRISS, in the most favourable case, that is, the target with the highest flux for which the detector does not saturate, it is possible to detect companions at a contrast of 10^{-4} at $1\lambda/D$.

Finally, in appendix, I show that the degradation of the contrast detection performance at low separation can be attributed to a degeneracy between the signature of a binary and a tip-tilt error on the pupil plane phase. Such an error corresponds to a displacement of the image on the focal plane. Because this signature is killed by the kernel transformation, and corresponds to a shift of the image, it can be considered to be a fundamental limit tot he detection of companions at very low separations ($> 0.25\lambda/D$).

Acknowledgements

Merci d'abord à Frantz Martinacche pour avoir accepté d'encadrer cette thèse. Je tiens également à remercier David Mary pour m'avoir introduit aux statistiques, Mamadou N'Diayé pour ses conseils sur tout le reste, à Romain qui a su supporter mes monologues interminables.

Je tiens également à remercier le jury, dont les remarques ont grandement contribué à la forme de ce manuscrit.

Contents

List of Figures

Introduction

The birth date of modern astronomy could be set at the first use of a telescope for night sky observations by Galileo around 1609. For the first time, astronomical observations where not limited to the power of the unaided eye. Although telescope design and performance evolved for the next two and a half centuries, these telescopes were still very much an extension of the human eye, and required an observer to peer at the sky through an eyepiece. This changed with the apparition of photography in the mid 19th century, which enabled observers to store and share objective reproductions of their observations. More recently, photographic plates were replaced by digital detectors, allowing for an advanced mathematical treatment of images. Although many other observing techniques have emerged, images remain a staple of observational astronomy to this day, and for good reasons: they are extremely information rich, and, as a direct extension of our sense of sight, are very easy to interpret and understand.

To analyse images, one can consider them to be a "perfect" representation of a scene, meaning the telescope itself is assumed to be a perfect image forming device, that merely projects a scene of the sky on a focal plane, where it is captured by a detector. This approximation can be valid in some circumstances, for instance if the field of view is very large, the size of the resolution elements can be neglected, and the scene dominates the obtained image. However, for high angular resolution observations, this approximation falls apart, as the effects imposed by the observatory come to dominate the image.

These effects can be broken into two main components: diffraction, and effects of optical defaults. Diffraction is inherent to the telescope, and is a function of the geometry (the shape and the size) of the entrance aperture. Any system is subjected to diffraction, which determine the highest image quality possible. The diffraction limit can only be improved upon by observing at shorter wavelengths, or by increasing the size of telescopes. As it is improved, the relative importance of other factors that limit the image quality, such as optical defaults increases.

For big ground based telescopes, the resolution attainable is limited not by diffraction, but by atmospheric effects. The atmosphere is in constant movement, with eddies forming at different scales. This movement can be easily observed in smoke of steam: a column of smoke, be it produced by a lit cigarette or a camp fire always ends up being broken up and dissipated by this turbulence. In astronomy, the fact the atmosphere limits resolution has been known for a long time, and limiting the amount of turbulence has been a major concern in siting observatories for centuries. Concepts of

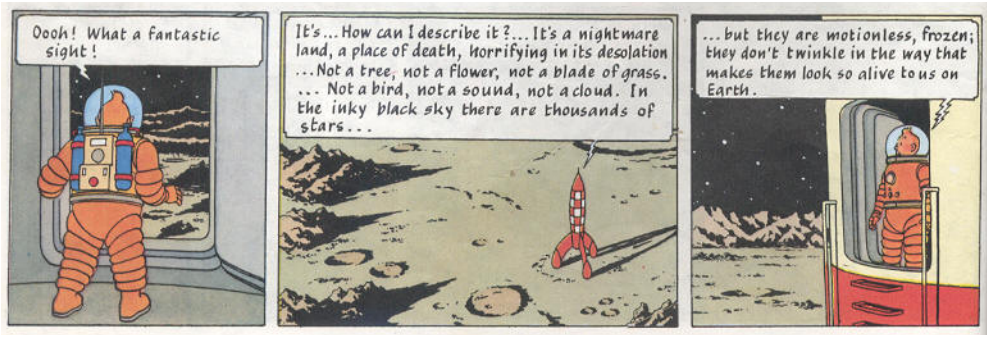

FIGURE 1: What Tintin notices here is the absence of twinkling of the stars seen from the surface of the Moon, which is an effect of atmospheric turbulence.

FIGURE 2: Images of NGC 2261 taken with the Hale telescope in 1949, and by the Hubble Space Telescope in 1999. Even though the Hale telescope is twice the size of Hubble, it produces less sharp images. The atmosphere explains this difference: the 1949 image is limited by atmospheric effects, while the 1999 Hubble image is diffraction limited.

space telescopes, for instance have been proposed as soon as leaving Earth's atmosphere has been possible, in the 1940s. In 1953, Hergé had Tintin note how stars seen from the surface of the Moon do not twinkle in the Explorers on the Moon album quoted in Fig. 1.

To overcome this limit, three approaches are commonly employed. The first one is to overcome the atmosphere by avoiding it altogether, using a space based observatory. This approach, employed for example by the space telescope Hubble have enabled a spectacular leap in resolution, as illustrated in Fig. 2. Other, less costly methods have also be employed. Historically, the first one to be employed may have been interferometry, which consists in simplifying the image so that the effects caused by an observed object can be separated by those caused by optical defaults. This term encompasses a family of techniques, which can enable to the recovery of more or less information on the observed object, with different degrees of robustness to optical defaults. Most recently, a technique have emerged that can correct

the atmospheric turbulence, adaptive optics, or AO. Although it has been experimented with since the 1980s, the past decades have seen the first time AO systems have become able to correct images well enough to restore the diffraction limit on images acquired with large, ground based telescopes.

With AO and space based observations, it becomes possible to adapt some techniques carried over from interferometry to analyse images, and to produce a set of interferometric observables that are robust to small optical defaults, called kernel phases. Kernel phases rely on modelling the propagation of optical defaults in a pupil plane to the phase of the complex visibility of the image.

My work during this thesis has focused on the extraction, and, most importantly on the statistical treatment of these observables. In this text, I focus on statistical detection of binaries using unsaturated images obtained in non-coronagraphic mode. This is especially useful at very close separations, where light leakage from coronagraphs become important, and the bright diffraction pattern of the central source tends to drown the image of surrounding structures.

Being quite new, the kernel method produces observables that have to be understood, and treated appropriately to produce useful scientific results. For once, the distribution of the errors that affect kernel phases must be known. Once this is done, systematic detection procedures can be employed to both determine whether an object is a binary, and its parameters.

I then show a practical application of these tests on kernel phases for the detection of companions to Y type brown dwarfs. Y dwarfs are little known object (with a small population of a few dozen objects detected to date). Determining their multiplicity rates could help better understand the physics of star formation to probe the bottom of the mass distribution. Observing them from the ground is however difficult because of their very low flux, especially in the near infrared. The James Webb Space telescope should alleviate these difficulties, and allow for much better observations of these objects. These detection procedures have proven to be an effective complement to the JWST coronagraphs, as they can reveal companions at a fraction of the separations accessible to them.

One of the techniques that exemplify this approach is interferometry. Interferometry is able to recover information about the spatial structure of objects, but does not use images per see. It relies on a combination of the light of several telescopes to recover very specific spatial information, not in the form of images, but of visibilities. The aspects of the spatial structure of the object, or visibilities the method recovers is set by the aperture configuration. This technique therefore comes with a degree of robustness by design. Furthermore, interferometry and imaging can be described with a common mathematical framework, enabling the adaptation of interferometry-specific methods to images.

This dissertation comports three chapters. Chapter 1 presents the problem of high resolution in Section 1.1 and introduces the interferometric concepts

and the kernel method in Section 1.2. Chapter 2 introduces the mathematical framework allowing the construction of kernel phases in the small aberration regime in Section 2.1, presents kernel phases in Section 2.2, and the errors that can affect them in Section 2.3. Chapter 3 presents the statistical tools I have developed to detect features using kernel phases, with an introduction to statistical detection in Section 3.1, Section 3.2 then introduces hypothesis testing, and present three different tests adapted to the detection of binaries. Section 3.3 examines a practical use case: the detection of binaries in JWST images of cool brown dwarfs, and Section 3.4 gives the expected detection performance for each of the proposed tests, taking real world Y dwarfs.

Chapter 1

Image formation, interferometry and the limits to resolution.

Modern astrophysics is a diverse and rich scientific discipline which attempts to understand the Universe by combining observations, experiments and models, and has been in many ways driven by technological progress. For thousands of years, the only direct observation tool available to astronomers was their own sense of sight. Compared to the tools available to astronomers in the twenty-first century, the unaided eye is a limited instrument: it can only detect a few thousand stars out of the hundred billion in the galaxy, is sensitive to only a tiny slice of the electromagnetic spectrum, and its resolution is too low to resolve the nearest planets, and distinguish them from distant stars. In this chapter, I present a brief overview of the physical processes that limit angular resolution and of the techniques that can improve it, with Section 1.1 explaining the concepts of resolution and how images are formed in Section 1.1.1 and how the wave nature of light naturally limits the resolution attainable by a telescope in Section 1.1.2. Section 1.2 shows how the problem of resolution can be reframed in the Fourier domain, and presents techniques exploiting the Fourier domain information: speckle interferometry, seen in Section 1.2.1, aperture masking interferometry, see in Section 1.2.2, kernel phase interferometry (the method we focus on in this dissertation) in Section 1.2.3, and long baseline interferometry in Section 1.2.4.

1.1 Image-forming telescopes and angular resolution

Four hundred years ago the invention of the telescope suddenly enhanced our observing capabilities. The first recorded use of this instrument for astronomical observations, by Galileo Galilei sparked a scientific revolution. In the span of a few weeks, Galileo discovered the phases of Venus, Sun spots, the rotation of the Sun and the satellites of Jupiter. These discoveries upturned cosmological models dating back to Antiquity, and propelled astronomy into the modern era. Since then, progress in astronomy have been closely linked with improvement of telescope design, and the worldwide community of astronomers is now looking forward to the completion of 30-metre class

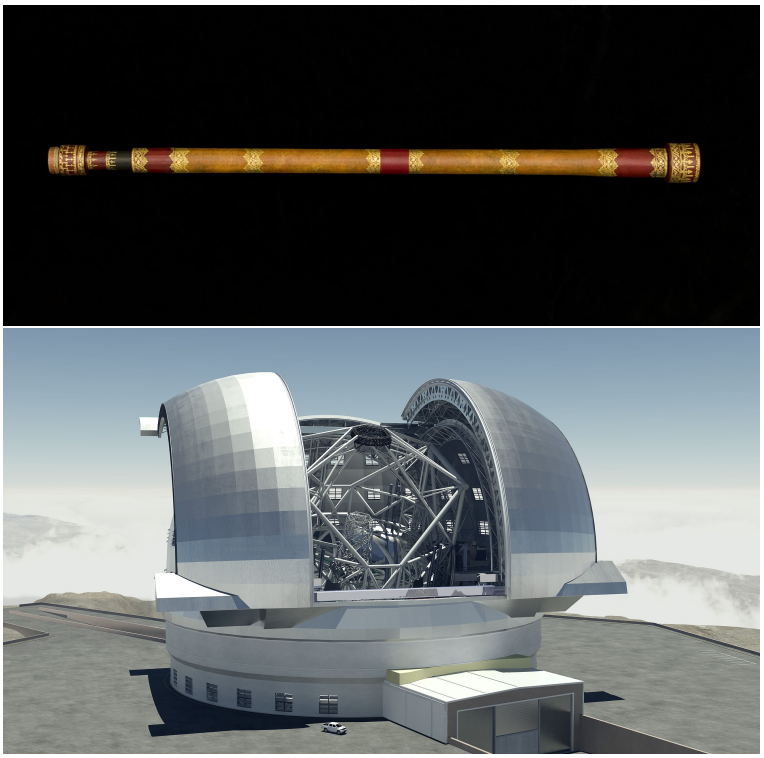

FIGURE 1.1: TOp: reproduction of the two-lens telescope used by Galileo in his observations of the moons of Jupiter, the phases of Venus and sun spots in 1609, *image courtesy of the Museum of Applied Arts & Sciences, Sidney, Australia*. Bottom: rendering of the Extremely Large Telescope(ELT), under construction at Cerro Armazones, Chile. *Image courtesy of Swinburne Astronomy Productions/ESO*

telescopes, such as the European Extremely Large Telescope (E-ELT) with its 39 metres wide primary mirror in Chile (McPherson et al., 2012), or the Thirty Meter Telescope (TMT) in Hawaii (Usuda et al., 2014). From the first aid to human vision to these next generation observatories, the collecting area has been multiplied by two million and the maximum angular resolution by 1400.

1.1.1 Image formation

Telescopes are image-forming devices: the information they provide is an estimation of the brightness distribution of the sky. Prior to the second part of the nineteenth century, images would ultimately be formed in the eye of the observer, who would be an integral part of the observing chain. With the advent of photography, images could be formed on photosensitive plates for later interpretation, allowing for the separation of the observation and the interpretation of its results. After being used for a century, these plates were replaced with electronic detectors, that dematerialise the information and allow for its near instantaneous sharing, and digital processing.

An essential property for an image is its angular resolution, or the smallest details it can help discerning, which governs the amount of spatial information that can be obtained while observing a given field of view. The resolution is expressed as the angular size of this smallest discernible feature. It is linked to the size of the image of a luminous point on the focal plane. In practice, this spot is never a point, and forms a pattern of a given size and shape, called the point spread function, or PSF. The size of the PSF in relation with the plate scale[1] can be used to approximate the maximum angular resolution afforded by a given telescope using Rayleigh's criterion, which states that the maximum angular resolution afforded by a telescope is determined by the diameter of the core of the PSF (Rayleigh, 1879). This principle is illustrated in Fig. 1.2, where it is apparent that the two point sources can be easily distinguished if they separated by more than a diameter of the core of the PSF.

For the most part, the design of telescopes relies on geometric optics, where light is described as a collection of rays propagating in straight lines, which are deviated either by reflection, where a ray hits a mirror and is reflected in a direction which depends on the orientation and shape of the mirror, or by refraction, where a ray crosses a boundary between two mediums of different refraction indices, with the shape of the surface and the refractive indices of the medium controlling the deflection of the ray. Lenses, mirrors or a combination of the two are arranged and shaped to guide the light from the pupil to a focal plane, where an image is formed.

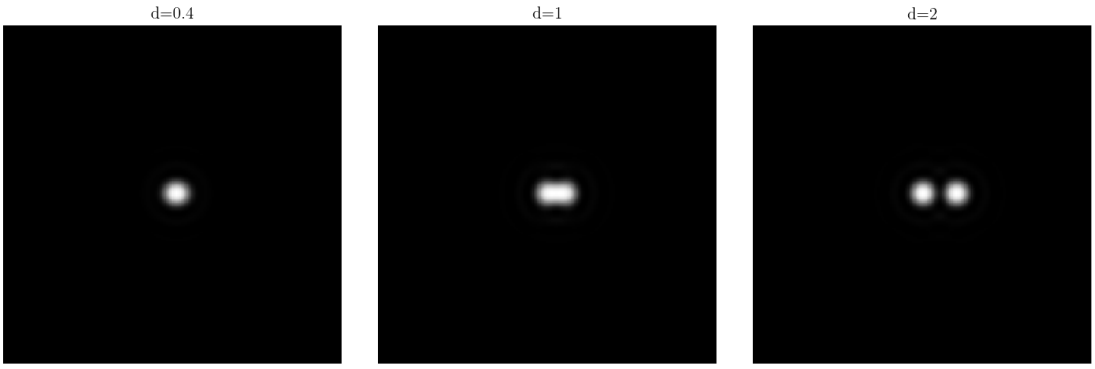

FIGURE 1.2: Rayleigh's resolution criterion states that two sources are resolved if their separations are at least equal to the distance between the maximum and the first minimum of the PSF, as is the case on the central image. Here, d,the angular distance between the two source points is expressed in half diameters of the core of the PSF.

[1]The plate scale links the angular size of a given feature on the sky to the physical size of this feature on an image formed on the focal plane.

In the geometric optics description, the resolving power is limited by the ability of the optics to focus the light from a point of a scene on a single point of the focal plane. This description does not exclude the possibility of a punctual PSF, and thus of infinite resolution, as long as magnification is high enough and the placement and construction of the optical system perfect. In reality, the extent of the PSF is constrained by the undulatory nature of light through diffraction.

1.1.2 Diffraction: a fundamental limit created by the wave nature of light

In the context of astronomical observations, diffraction was predicted by Fresnel (1823), and first observed by Herschel (1828), who conducted a series of experiments where he would observe a bright star with a telescope, and change the diameter of the entrance aperture. In the purely geometrical optics description, reducing the size of the aperture improves the optical quality of the instrument, as it diminishes the impact of optical defects. On the contrary, Herschel observed that:

> *When the aperture was much reduced (as to one inch, for a telescope of 7 feet focal length) the spurious disc was enlarged to a planetary appearance, being well defined, and surrounded by one ring only, strong enough to be clearly perceived, and faintly tinged with colour.*

Two things are of note there: firstly, the size of the diffraction pattern (called the "spurious disk“ in the quoted passage) increases as the diameter of the telescope aperture is decreased. Secondly, the coloration of the rings indicates that diffraction is wavelength dependent.

Diffraction can be explained using the Huygens-Fresnel principle, where each point of the entrance aperture of the telescope is treated as a coherent emitter. Aberrations, and other optical defaults can be described as affecting the phase and amplitude of these emitters, mainly by imposing phase delays on these virtual emitters, depending on their position on the pupil. An optical apparatus that can form an image works by placing the focal plane at the optical infinity of the aperture. In this case, the complex amplitude of the light on the focal plane is the coherent summations of the complex amplitudes across the whole entrance aperture, as the light originating from each point in the aperture interferes with the light originating from every other point. The phenomenon of interference is illustrated in Fig. 1.3, for two emitters producing waves in water. The same fundamental principles apply for every other wave, including the electromagnetic fluctuations making up light.

Summing the complex amplitudes of the electric field across the entrance aperture is equivalent to taking the Fourier transform of the complex intensity across the aperture: the electric field at each point of the aperture is considered

to oscillate at a given frequency, amplitude and phase.[2] During an exposure to form an image, this amplitude is averaged over a long timescale compared to the typical oscillation period of the electric field that makes up light. Thus, the image on a focal plane placed at infinity compared to the pupil can be written

$$\mathcal{I} = |\mathcal{F}(\mathcal{P})|^2 \tag{1.1}$$

with $\mathcal{P}$ the complex amplitude across the pupil.

The diffraction pattern produced by light after passing through a circular, unobstructed aperture is called an Airy disk. After being observed by Herschel, its shape was mathematically determined for a telescope with a circular aperture by Airy (1835). The Airy disk comports rings of decreasing intensity as the distance from the brightest part of the image increases. For monochromatic light, the intensity of the darker parts of the rings goes down to zero. The size of the central patch, defined as the radius of the ring formed by the first zero of Airy's function is given by the relation

$$R = 1.22\frac{\lambda}{D}, \tag{1.2}$$

where D is the diameter of the entrance aperture, and λ the wavelength at which the observation is made. The description of rings tinged with colour made by Herschel corresponds to a superposition of the diffraction patterns created for all the wavelengths that constitute white light.

Although daily life is replete with overlapping light sources, we do not usually observe interference fringes. The ability of two fields to interfere is estimated by a quantity called the mutual coherence: to produce sustained interference patterns, two light sources need to be mutually coherent. If the waveform of either of both source changes rapidly, the interference pattern they produce also changes at a speed similar to the frequency of the incoming waves. All light sources, except for very high quality lasers are mutually incoherent, because light is emitted by semi random processes at the atomic level. Since there is not physical process synchronising the emission of two point on the filament of a light bulb or in the photosphere of a star, every point in any given light source emits light that is mutually incoherent with the light emitted at any other point. The only way to produce interference with light is therefore to have it interfere with *itself*, by splitting it into different beams that are later recombined, with path difference within the coherence length. The coherence depends strongly on the nature of the source. with lasers achieving several centimetres in the optical domain. For natural, polychromatic sources, the coherence length is governed by the width of the considered spectral band $\delta\lambda$, and is given by $\lambda^2/\delta\lambda$.

[2] Polarisation is another property of light, stemming from the directionality of the electric field. It has many relevant applications in astrophysics, since it can help constrain some physical and chemical properties of sources, but is not discussed in this manuscript, where all light is considered to be non-polarised.

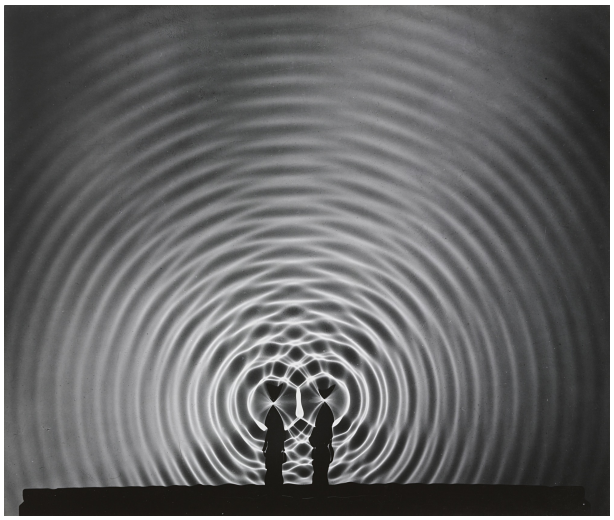

FIGURE 1.3: Photography illustrating the phenomenon of interference on the surface of water. Two sources of excitation produce two systems of waves. When waves meet, they interfere, either constructively (amplifying oscillation) or destructively (the waves cancel each other). In this case, the interference results in the formation of calmer zones, which seem to radiate outward from the sources of the oscillation. *Berenice Abbott, Interference of Waves, 1958-61, image courtesy of MoMA*

As a result of diffraction, a point on a scene produces one PSF, centred on a position a focal plane, linearly dependent on the position of the said point on the scene.[3] Mathematically, this translates as the image being the result of the convolution of the spatial intensity distribution making up the scene and of the PSF

$$\mathcal{I} = \mathcal{O} \circledast PSF. \tag{1.3}$$

This relation means the resolution attainable with a telescope is limited by the knowledge of the PSF, as well as by its size (a smaller PSF yields a higher resolution). The extent of a PSF dominated by diffraction is directly proportional to the inverse of the size of the entrance aperture, but the shape of said aperture also plays a crucial role.

1.2 Resolution in the Fourier domain

The way the PSF and the object combine to form an image can be made more explicit by considering the Fourier transform of the image, as the Fourier transform of the convolution of two functions is the product of the Fourier transforms of each function. Thus, the relation Eq.

[3] This thesis will not get into the cases where the PSF changes depending on the position of source of light, such as for fields of view greater than the isoplanatism angle, or coronagraphic observations.

eqrefeq:convol can be transformed into

$$\mathcal{F}(\mathcal{I}) = \mathcal{F}(\mathcal{O}) \times \mathcal{F}(PSF). \tag{1.4}$$

$\mathcal{F}(\mathcal{I})$ is the complex visibility measured by the telescope, and is also noted $\mathcal{V}$. $\mathcal{F}(\mathcal{O})$ is the complex visibility of the object, the information an observer attempts to recover, also noted $\mathcal{V}_0$. $\mathcal{F}(PSF)$ is the optical transfer function, or OTF of the telescope, and encapsulates all the effects caused by the telescope and by the atmosphere.

In the Fourier domain, spatial information is described not in terms of geometrical position, but as a collection of spatial frequencies. Higher spatial frequencies are associated with finer details, and lower spatial frequencies with larger structures.

The amplitude of the OTF of a telescope is a gain on the complex visibility of the object, for each spatial frequency. For spatial frequencies higher than D/λ, the gain is zero: all spatial frequencies higher than this limit are filtered. This is the true limit to spatial resolution. This limit does not necessarily preclude the observation of features smaller than λ/D, provided they are associated with a complex visibility that exists at spatial frequencies smaller than D/λ. This recovery is easiest for objects with a structure that can be easily parameterised, such as a disk or binary. Thus, in the Fourier domain, there is no hard resolution limit *per se*, but the recovery of spatial information for very small features becomes more and more limited by the noise on the complex visibility associated with the image, as the amplitude of the complex visibility of a smaller spatial feature becomes smaller at smaller spatial frequencies. The practical resolution is limited by the signal to noise ratio for the measurement of the complex visibility at the spatial frequencies sampled by the aperture.

For centuries, the resolution attainable with telescopes was not limited by diffraction, but by the atmosphere. It is in a state of constant agitation, as it constantly seeks thermal equilibrium. This creates pockets of air at slightly different temperatures, and therefore different refraction indices. Thus, the light reaching different points of the entrance aperture is affected by different delays, causing the PSF to deviate from what it would be in the absence of atmospheric effects. These delays are random, but they exhibit correlations at small spatial scales. The distance over which two points in the aperture experience similar delays is approximated by the Fried parameter r_0 (Fried, 1966).

1.2.1 Speckle Interferometry

The atmosphere evolves on timescales typically much shorter than the exposure time usually required for useful observations, with the coherence time of the atmosphere being measured in milliseconds, and most astronomical observations relying on exposures seconds or longer. For instance, at $\lambda = 500nm$, Lawrence et al. (2004) reports coherence time values ranging from 2.9 ms at Mauna Kea to 7.9 ms on Dome C in Antarctica. Because of

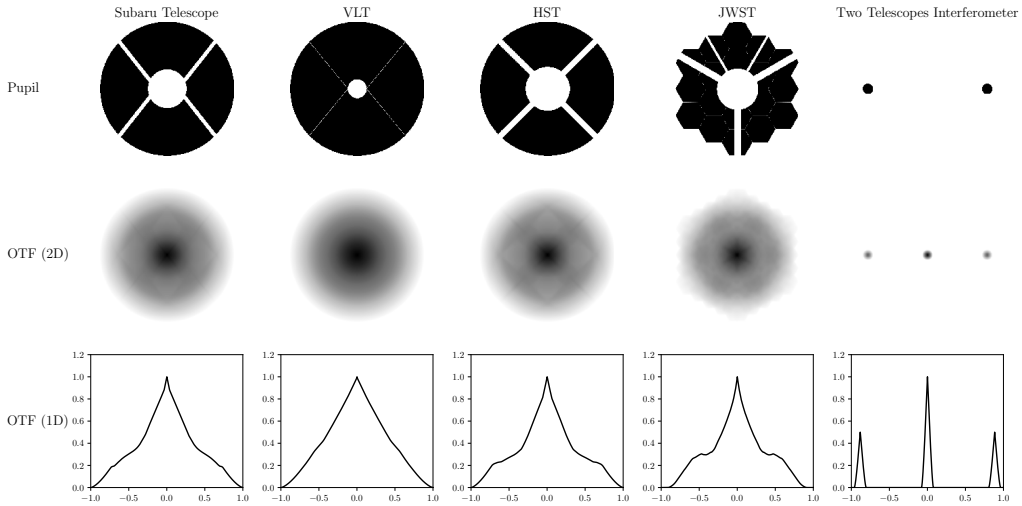

FIGURE 1.4: The optical transfer functions for the Subaru telescope, the Very Large Telescope, The Hubble Space telescope and the James Webb Space Telescope. The diameters of all of these apertures is arbitrarily taken at 1. At spatial frequencies larger than D/λ, the OTF drops to zero, giving a resolution limit close to Rayleigh's criterion (although spatial information can be recovered beyond this limit). For a aperture made up of a pair of holes, or pupils, there is a gap in the coverage of some spatial frequencies, while the OTFs for full pupil telescopes allow the sampling of every spatial frequency up to the cutoff.

the long exposures compared to the coherence time, the typical observation through the turbulent atmosphere is an incoherent superposition involving all the fluctuation of the turbulence throughout the exposure. On the focal plane, the long exposure PSF is a seeing spot, of a size λ/r_0, with a loss of the spatial information for smaller features.

This loss can be explained in the Fourier domain. Phase delays in the aperture affect mostly the phase of the OTF, and Fig. 1.4 shows how a spatial frequency can be linked to a pair of points in the pupil. For spatial frequencies corresponding to an ensemble of points farther apart than the Fried parameter, the phases differences are uncorrelated, and thus the modulus of OTF for these spatial frequencies tends towards zero for long exposures. For smaller spatial frequencies, the OTF is non zero thanks to the correlation in the fluctuations of the phase. For long exposures, atmospheric turbulence therefore causes higher spatial frequencies to be irremediably lost. On the focal plane, this loss translates into the formation of a seeing spot. The size of the seeing spot depends on the weather, the wavelength and the observation site, but rarely dips under the arcsecond in the visible. This means that resolution wise, in the visible, increasing the diameter of the primary mirror of a telescope beyond 30cm or so bears no advantage resolution-wise if no technique is employed to deal with the effects of atmospheric turbulence.

One way to counter the adverse effects of atmospheric turbulence is to use exposure times shorter than the time of coherence. The short exposure images obtained in this case contain features of a size λ/D, called speckles (Goodman, 1986): small angular features are not so much lost as they are scrambled. These speckles are dispersed in the focal plane, and are the brightest in a region of the size of the seeing spot. The evolution of the atmospheric turbulence causes these speckles to move around randomly on the timescale of the coherence time of the atmosphere. In terms of Fourier domain information, for short exposures, the averaging of the OTF does not occur. Instead, the random phase fluctuations are frozen in time, and the amplitude of the OTF does not average to zero for spatial frequencies higher than r_0/λ. By summing the squares of the observed complex visibilities of many short exposure images, the modulus of the complex visibility of an observed object can be recovered at spatial frequencies higher than the long exposure limit of $r_0\lambda$, although at the cost of an attenuation of the modulus of the recovered complex visibilities for high spatial frequencies compared to the turbulence free case. This technique is known as speckle interferometry (Labeyrie, 1970), and was the first one to allow recovery of spatial information beyond the seeing limit for full pupil telescopes.

if speckle interferometry allows for the reconstruction of the amplitude of the complex visibility of an object, squaring the Fourier transforms of the short exposure images intrinsically destroys the phase information of the complex visibility. The phase fluctuations created by atmospheric turbulence also badly scramble the phase of the measured complex visibility, as these fluctuations are very important (the optical path differences, or OPDs across

a large pupil can be several wavelengths). In some circumstances, this phase noise can be limited by understanding how the phase noise propagates from random OPDs across the pupil to the phase of the measured complex visibility. As shown in the rightmost pupil in Fig. 1.4, a pupil composed of two holes can be considered to sample only one spatial frequency, which depends on the relative positions of the holes. The vector that links the two holes in the plane of the aperture, and by extension the unique spatial frequency this pair can be said to sample is called a *baseline*. The phase of the OTF for this spatial frequency is determined by the OPD between the two holes. In the case of a full pupil, this description also holds, but then the OTF for given spatial frequency is the result of the sums of the OTFs of every possible pair of points in the aperture that sample said spatial frequency. For this reason, a full pupil is called redundant: each spatial frequency is sampled by a large number of pairs of points across the entrance aperture. Indirect measurements of the phase of the complex visibility are also possible by using the correlations of the Fourier transforms of images (Lohmann, Weigelt, and Wirnitzer, 1983; Ghez et al., 1990).

1.2.2 Aperture masking interferometry: turbulence-proofing a telescope

Redundancy makes the phase information of the complex visibility of an object difficult to recover, since the phase of the OTF for a redundant aperture at a given spatial frequency is a non linear combination of the phases shifts imposed by turbulence for every pair of points making up the entrance aperture. In the absence of corrections for the effects of turbulence, the measurement of complex visibilities can be improved by mitigating redundancy in the aperture. This is done through a technique called aperture masking interferometry, where a non-redundant mask (NRM) composed of an array of holes such as each pair of holes samples a single spatial frequency, with each hole smaller than r_0 is placed in a pupil plane, to alter the aperture's geometry. This technique was originally developed in the early 19th century (see Tuthill (2012) for an overview of the history of the technique) and used by Stephan (1874), and to make the first stellar diameters measurements.

The non-redundant nature of the aperture affects the recovery of both the modulus and the phase of the complex visibility of the observed object. The amplitude of the OTF of a non-redundant aperture is independent from its phase, thus the amplitude of the complex visibility of an object can in principle be recovered from a single short (below the coherence time of the atmospheric turbulence) exposure image, as shown in Fig. 1.5. The phase information can also be recovered using the closure phases technique, which exploits the fact there exists a linear relation between the phase of the OTF for each spatial frequency and the phase shifts imposed by turbulence for each of the holes making up the non-redundant aperture (Jennison, 1958). Therefore, non-redundant interferometry enables the recovery of the full complex visibility

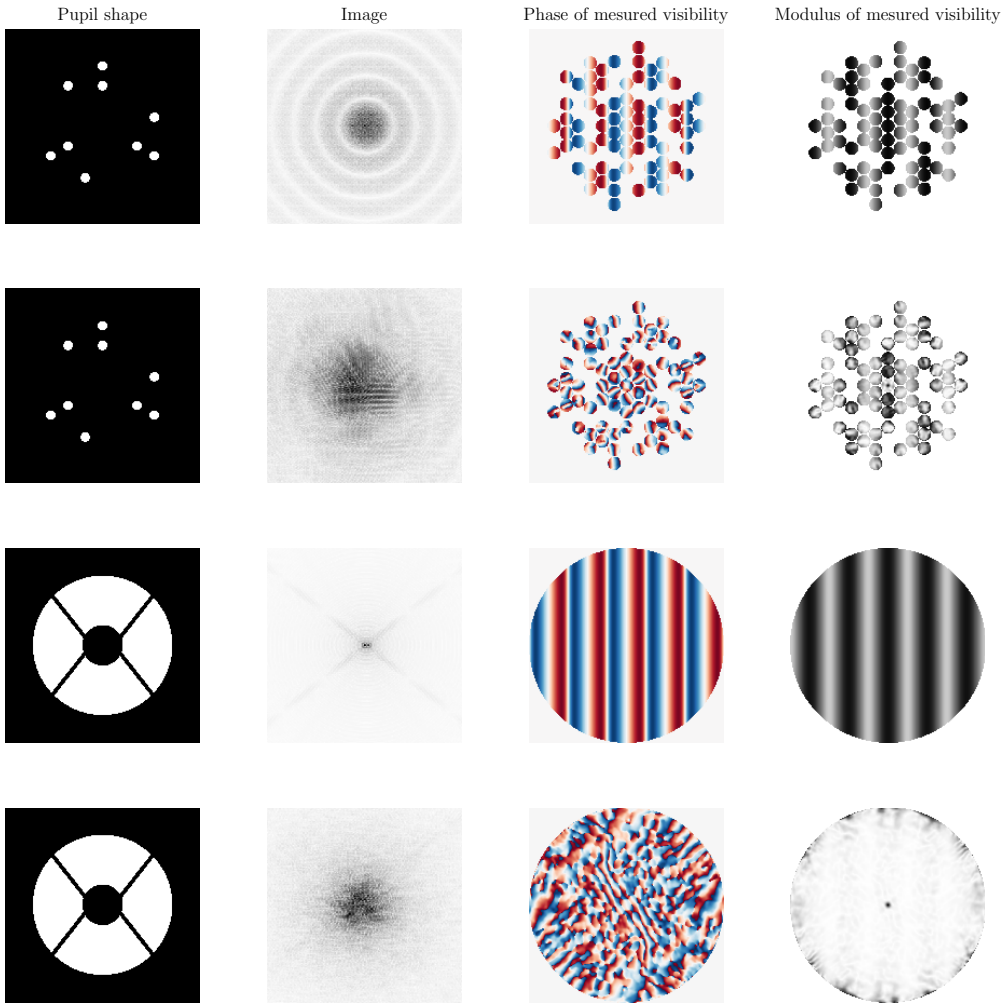

FIGURE 1.5: The image and the associated complex visibility obtained using a non-redundant aperture mask (NRM) and the full pupil of a telescope, in the absence of aberrations (first and third rows), and through atmospheric turbulence (second and fourth rows). In both cases, the observed object is a double star. For a redundant pupil, the object visibility ends up badly scrambled by aberration, both in phase and in modulus, while a non-redundant configuration preserves the modulus better. On the downside,the non-redundant configuration occludes most of the area of the pupil, severely limiting transmission.

information (Nakajima et al., 1989) by combining closure phases, with the measurements of amplitudes.

1.2.3 Kernel phase interferometry

The issue of atmospheric perturbations the techniques of speckle interferometry, non-redundant aperture interferometry and bispectral analysis originally set out to solve have been largely overcome by technical progress in the correction, or even avoidance altogether of atmospheric effects *via* adaptive optics and space telescopes. For high angular resolution imaging, it is now possible to obtain "diffraction limited" (images close to the diffractive theoretical resolution limit) from 8 meter class telescopes thanks to adaptive optic (AO) systems, such as the Gemini Planet Imager on the Gemini South Telescope (Macintosh et al., 2006), SPHERE on the Very Large telescope (Dohlen et al., 2006) or SCExAO on the Subaru Telescope (Lozi et al., 2018). In space, the upcoming James Webb Space Telescope (JWST) will rival current ground based 8 metre class telescopes with a 6.2 metres primary mirror. However, although telescopes and corrections systems have progressed to the point that the OPD is smaller than a wavelength across the aperture, these errors have not been brought down to zero, and the uncertainty these residual OPD bring are still significant to the recovery of high angular resolution information. In particular, residuals are still a large component of the error budget for high contrast imaging, where the object is dominated by a central star, and thus its complex visibility has a amplitude and close to 1 and a phase close to 0, and even a small noise from the OPD can be significant. Aperture masking interferometry thus remains a popular observation method in the age of extreme adaptive optics and space telescopes, thanks to the self calibration of the phases of the visibilities made possible by closure phases.[4] Gauchet et al. (2016) shows how NRM observations on SPHERE can complement coronagraphic observations to probe separations below λ/D while providing contrasts of the order 10^{-3} on a variety of stellar disks, JWST will also be capable of making NRM observations (Sivaramakrishnan et al., 2012) to probe small separations, although, as I will show in the latter part of Chapter 3, it is also possible to obtain compelling performance without the use of a mask.

For AO corrected of space telescope images, the OPDs across the aperture are generally less than a wavelength. In this regime, redundant configurations no longer suffer the loss of a linear relation between the phase of the OTF and

[4] In the case of adaptive optics, correction residuals can reach several hundreds of nanometres of OPD (Kooten, Doelman, and Kenworthy, 2020), which corresponds to a significant fraction of the wavelength for observations in the near infrared. In the image, these residuals create faint speckles, which are still significant compared to the typical flux of many companions, especially behind a coronagraph which filters out most of the on-axis light of the star, but do not affect the speckles. For high contrast applications, the contribution of high performance AO is invaluable, but self calibration, such as provided by closure phases with NRM interferometry is still desirable.

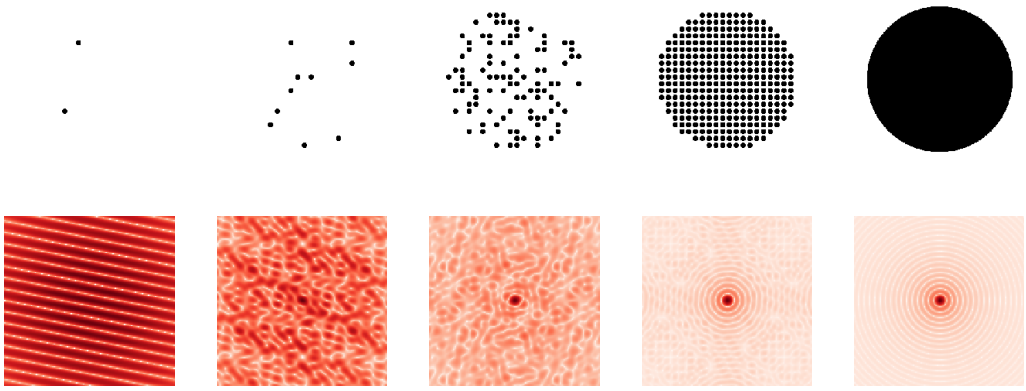

FIGURE 1.6: Top: entrance pupil, composed of either a collection of holes, or of a continuous entrance pupil. Bottom: interference patterns for these arrangements. As the number of holes, arranged on a disk increase, the interference pattern they form gets closer and closer to the diffraction pattern of a circular aperture. This is an ideal case, shown for monochromatic light with a coherence length much greater than the maximum path difference.

the OPDs. Furthermore, this linear relation can be used on full pupil images. This is done thanks to a discrete model as the aperture, as the image produced by a a telescope with a given aperture geometry can be approximated as the interference figure produced by an array of pupils places in a way that emulates the aperture's geometry, as shown in Fig. 1.6. The phase of the OPD can then be modelled by a set of linear relations between the OPDs over each of the virtual pupils used to represent the aperture, and linear combinations of the phases of the complex visibilities sampled by the array of virtual pupils can be constructed from these relation to produce observables that are robust to the OPD over the aperture. These observables are called *kernel phases* (Martinache, 2010), the main subject of this dissertation. Kernel phases do not require the introduction of a non-redundant mask, allowing for more flux to reach the detector (a NRM typically cuts over 80% of the incoming flux). They also permit the exploitation of the full Fourier domain. Their construction is explained in more details in the first part of the next chapter, and Chapter 3 presents methods for the statistical treatment of these observables, and constitutes the core of my contribution to the development of this method.

1.2.4 Long baseline optical interferometry

On another front, optical interferometry has also emerged as a technique able to provide even higher resolution than either imaging or NRM interferometry. This approach has been used for scientific observation as early as 1813

with Thomas Young's measurement of the diameter of human red blood cells (Young, 1823). In astronomy, the first attempt at an interferometric observation was made by Édouard Stephan (Stephan, 1874), who, drawing from the work of Fizeau (1868) masked the aperture of a telescope to overcome the effects of atmospheric turbulence, performing the first documented astronomical NRM interferometry observation. The potential of interferometry in the improvement of the maximum attainable resolution lies however in the possibility of extending the lengths of baselines beyond the limit of the diameter of the aperture of a single telescope. The first step in this direction was taken by Michelson and Pease (1921), who used a beam across the entrance of a telescope on which two small mirrors would slide to augment its effective diameters. Over half a century later, Labeyrie (1975) achieved the first successful recombination of the light of several telescopes,[5] and opened the way for long baseline optical interferometry, which provides resolutions well beyond what current single mirror telescopes can achieve. Current optical interferometric arrays can sample baselines hundreds of metres long. Matisse for instance (Lagarde et al., 2006) can recombine the light collected by the four 8 metre telescopes at the VLTI, offering a simultaneous measurement of the visibilities at 12 distinct spatial frequencies. The CHARA array (McAlister et al., 1995) can form baselines up to 330m long. These arrays provide angular resolutions in the optical domain way greater than any individual telescopes, but at the price of a sparse coverage of the Fourier domain, which does not preclude the reconstruction of images from their observations (e.g Zhao et al. (2008)). The Hypertelescope concept aims at alleviating these issues by simultaneously recombining the light collected by a multitude of mirrors, to make up a massive diluted telescope (Labeyrie et al., 2001).

[5] I leave aside the early experiments in intensity interferometry (Hanbury Brown and Twiss, 1956) in this overview.

Chapter 2

Kernel-phases: aberration robust observables

The previous chapter presented an overview of different techniques aiming at improving the resolution of astronomical observation. This thesis concerns itself with the statistical treatment of kernel phases, constructed using images acquired with apertures of an arbitrary geometry with small OPDs, such as are obtained from ground based telescopes behind AO or space telescopes. In this chapter, I explain the mathematical foundations that enable the construction of kernel phases from such images.

I start by explaining the relationship between OPDs across pupils and the complex visibilities they sample when placed to form various arrays in Sec. 2.1, going up from a simple two pupils array in Sec. 2.1.1 to show how the phase of the complex visibility of an object suffers from a fundamental degeneracy with the shifts across each pupils. In Sec. 2.1.2, I explain how a third pupil can help lit this degeneracy by allowing for the construction of closure phases, and in Sec. 2.1.3, I show how for an array comprising four pupils that sample the Fourier domain redundantly, there exists a regime in which observables analogue to closure phases can be constructed from such an array. Sec. 2.2 introduces kernel phases proper, where linear algebra is used to construct observables robust to small OPDs across the aperture from any redundant pupils array, and Sec. 2.2.2 briefly touches on the fact such observables can also be obtained from the amplitudes of complex visibilities extracted from an image. The final section, Sec. 2.3 presents the errors that affect kernel phases, with Sec. 2.3.1 focusing on systematic errors, that stem from kernel phase residuals, and Sec. 2.3.2 focusing on how noises at the pixel level in the image propagate to kernel phases and on their modelling.

2.1 Interferometric visibilities and the impact of OPDs

The fundamental relation the constriction of kernel phases stems from the rearrangement of the convolution relation linking the PSF, the image and the intensity distribution of an observed object *via* a Fourier transform, such as,

as already written in Eq. (1.4), (1.4):

$$\mathcal{F}(\mathcal{I}) = \mathcal{F}(\mathcal{O}) \times \mathcal{F}(PSF). \tag{2.1}$$

This allows to draw a direct equivalence between the object and it's complex visibility, with

$$\mathcal{V}_0 := \mathcal{F}(\mathcal{O}), \tag{2.2}$$

the PSF and the optical transfer function or OTF of the optical system,

$$OTF := \mathcal{F}(PSF), \tag{2.3}$$

and finally the image and a measured complex visibility

$$\mathcal{V} := \mathcal{F}(\mathcal{I}). \tag{2.4}$$

Thus, as long as the convolution relation holds (i.e.for small enough fields of view and distant objects), the formation of an image on the focal plane can be considered equivalent to the measurement of a complex visibility. This relation is quite direct for sparse interferometers, where the "PSF" is a set of fringes. The amplitude and phase of the complex visibility is accessible as the contrast and positions of these fringes respectively, but it also holds in the general case, with PSFs resulting from the diffraction through a continuous aperture rather than from the interference of light from a pair of holes.

The OTF decomposes into its phase and amplitude; called the modulation transfer function (MTF) and the phase transfer function (PTF), which modulate the phase and amplitude of the measured visibility, such as

$$OTF := MTF e^{iPTF}. \tag{2.5}$$

For ideal observing conditions, if the OPD across the aperture are zero, the PTF is zero, and the MTF is the autocorrelation of the aperture. In the presence of an non zero OPD, the PTF is non zero, and the MTF is lower for high spatial frequencies for a redundant aperture.

2.1.1 Two pupil array

From an interferometric standpoint, the simplest aperture comports two distinct holes, or pupils of the same size. In this case, at the condition that the variations in OPD across each of the pupil is small, and that the size of the pupils is small relative to their separation, the aperture can be considered to sample a unique spatial frequency, which correspond to the relative positions of the holes. The longer the distance between the holes, the higher the spatial frequency sampled, and hence the spatial resolution (keeping the SNR on the measured complex visibility constant).

In Fig. 2.1, I show the Fourier plane sampling of the most basic interferometer, composed of two distinct pupils. This interferometer samples a

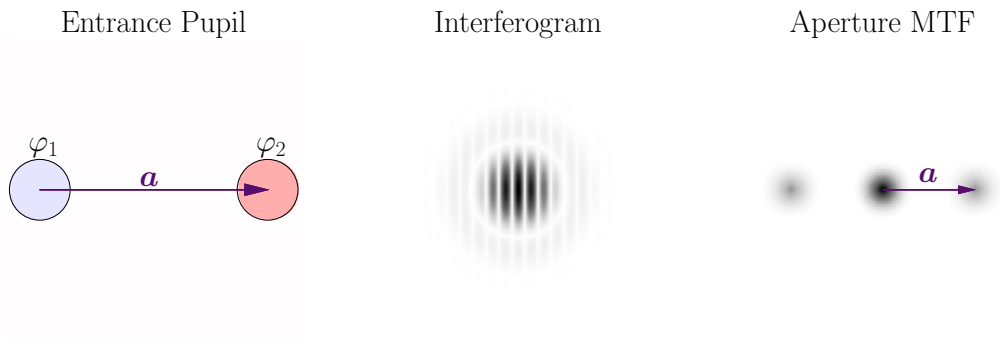

FIGURE 2.1: This aperture composed of two pupils can be considered to sample a single spatial frequency, which depends on the positions of the two pupils.

single baseline $\boldsymbol{a}$, which corresponds to the vector linking the two pupils. The interferogram produced by this disposition is a system of fringes, seen here modulated by the diffraction pattern created by the circular pupils.

Throughout this dissertation, vectors and matrices will be used extensively. Here is the notation convention I use: a lowercase of upper case letter in regular typeface, such as a or A represents a real or a complex scalar. A lowercase letter in a bold typeface, such as "$\boldsymbol{a}$" represents a vector. Finally, a bold, uppercase letter, like $\boldsymbol{A}$ represents a matrix.

The visibility $\mathcal{V}(\boldsymbol{a})$ measured by this interferometer is obtained from Eq. (2.1). Thus,

$$\mathcal{V}(\boldsymbol{a}) = \mathcal{V}_0(\boldsymbol{a}).OTF(\boldsymbol{a}). \tag{2.6}$$

The OTF for $\boldsymbol{a}$ is a result of the surface area of the two pupils and of the OPD across them. By convention, $MTF(\boldsymbol{a})$ can be set to 1 (in practice, the amplitudes of observed visibilities are calibrated from a measurement of the MTF, obtained by observing a point source). $PTF(\boldsymbol{a})$ is the result of the difference of the phase shifts on each pupil,

$$PTF(\boldsymbol{a}) = \varphi_1 - \varphi_2. \tag{2.7}$$

Thus, the complex visibility measured by this interferometer can be written

$$\mathcal{V}(\boldsymbol{a}) = \mathcal{V}_0(\boldsymbol{a})e^{i(\varphi_1 - \varphi 2)}. \tag{2.8}$$

By separating the complex visibility into its phase and amplitude, $\mathcal{V} := ae^{i\phi}$, we can independently link the phase and amplitude of the measured complex visibility to those of the complex visibility of the observed object:

$$\begin{aligned} a(\boldsymbol{b}) &= a_0(\boldsymbol{b}) \\ \phi(\boldsymbol{b}) &= \phi_0(\boldsymbol{b}) + \varphi_1 - \varphi_2 \end{aligned} \tag{2.9}$$

Therefore, the amplitude of the complex visibility of the observed object is

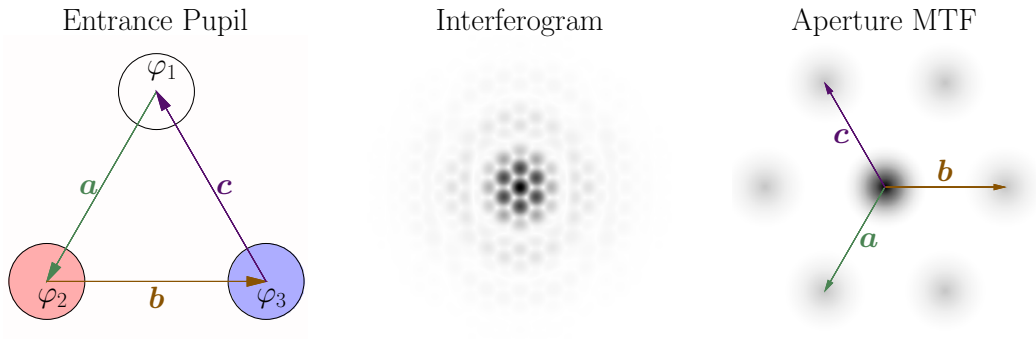

FIGURE 2.2: Example of the diffraction pattern and the Fourier domain coverage of a three holes interferometer. Here, we can see how each spatial frequency sampled by this interferometer corresponds to a baseline, or a pair of pupils.

directly accessible, but its phase is directly linked to the OPD, or phase shifts across each of the two pupils. Theoretically, this effect on the phase could be calibrated, but is rendered very difficult in practice because of the fast evolution of the OPDs. For a ground based observation, one would have to switch every few milliseconds to a calibration source and the scientific source, without relying on the position of the fringes formed by the observed object to produce a reliable measurement of the phase of a complex visibility from a two pupils interferometer.[1]

The phase noise corresponds to the *differential* phase shift between the pupils. Thus, having the phase shift φ_1 on the left hand side pupil in Fig. 2.1 and φ_2 on the right hand side pupil is, from the standpoint of the complex visibility measured by the baseline $\boldsymbol{a}$, strictly equivalent to having the phase shift $\varphi_1 - \varphi_2$ on the left hand side pupil, and no phase shift on the right hand side pupil. This property generalises to any interferometric array: it is always possible to define a reference pupil with no phase shift.

To obtain a reliable measurement of the phase, it is necessary to self calibrate it, by using only the visibilities measured from the observation itself rather than relying on calibration made using a separate source. This can be done using closure phases, which involve three or more separate baselines.

2.1.2 Closure phases from a three pupil array

By adding a pupil to the two pupil interferometer, one can sample up to three baselines. Since only the relative optical path difference governs the error on the phase of the complex visibility, only two parameters govern the measured phases of the complex visibilities. This is what makes the construction of a

[1] Current ground based interferometers do rely on the position of the fringes to stabilise them and obtain reliable interferometric measurements, using a *fringe tracker*. This system is the interferometric equivalent of adaptive optics. It improves dramatically the quality of these observations, but destroys the absolute phase signal.

single observable that depends in principle solely on the phase of the complex visibility of the observed object possible.

Let us consider the array composed of three pupils represented in Fig. 2.2. Each spatial frequency corresponds to a vector linking two pupils, or baseline: $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$. The phases of the complex visibilities sampled by this array are written:

$$\begin{aligned} \phi_{\boldsymbol{a}} &= \phi_0(\boldsymbol{a}) + \varphi_1 - \varphi_2 \\ \phi_{\boldsymbol{b}} &= \phi_0(\boldsymbol{b}) + \varphi_2 - \varphi_3 \\ \phi_{\boldsymbol{c}} &= \phi_0(\boldsymbol{c}) + \varphi_3 - \varphi_1. \end{aligned} \tag{2.10}$$

with φ_1, φ_2 and φ_3 the phase shifts caused by the OPDs for each of the pupils.

Since a global phase shift that affects all of the pupils has neither an influence on the PSF nor on the complex visibilities measured, only the relative optical path difference between the pupils counts. We can therefore arbitrarily set $\varphi_1 := 0$, and thus get

$$\begin{aligned} \phi_{\boldsymbol{a}} &= \phi_0(\boldsymbol{a}) + -\varphi_2 \\ \phi_{\boldsymbol{b}} &= \phi_0(\boldsymbol{b}) + \varphi_2 - \varphi_3 \\ \phi_{\boldsymbol{c}} &= \phi_0(\boldsymbol{c}) + \varphi_3. \end{aligned} \tag{2.11}$$

Then, there is a linear combination of $\phi_{\boldsymbol{a}}$,$\phi_{\boldsymbol{b}}$ and $\phi_{\boldsymbol{c}}$ that cancels out the terms due to the optical path difference:

$$\begin{aligned} \phi_{\boldsymbol{b}} + \phi_{\boldsymbol{a}} + \phi_{\boldsymbol{c}} &= \phi_0(\boldsymbol{b}) + \varphi_2 - \varphi_3 + \phi_0(\boldsymbol{a}) + -\varphi_2 + \phi_0(\boldsymbol{c}) + \varphi_3 \\ &= \phi_0(\boldsymbol{a}) + \phi_0(\boldsymbol{b}) + \phi_0(\boldsymbol{c}). \end{aligned} \tag{2.12}$$

This quantity is called a closure phase. One closure phase can be constructed for every triangle linking 3 sub-apertures. Practical uses employ pupil geometries with more that 3 sub-apertures (typical numbers range from 9 to 17), to maximise the number of closure phases that can be constructed and the coverage of the Fourier domain. On telescopes, the desired aperture geometry is obtained thanks to a non redundant mask, or NRM, placed on an intermediary pupil plane. The technique that consists in recovering complex visibility information, including closure phases from images obtained with such masked apertures is called NRM interferometry.

Since closure phases are obtained from the *phases* of the complex visibilities extracted from an image acquired behind a NRM, only the features that create such a phase signature are evident in these observables. This means only certain aspects of the structure of the observed object are revealed. Indeed, the phase of the Fourier transform of on sky flux distribution is insensitive to symmetrical features, which are encoded in its modulus. For instance, the phase of the complex visibility of a centrosymmetric disk is zero: the asymmetrical structures are made evident by the phase of complex visibilities. Any centrosymmetric object has a null phase signature.

Together with the visibility modulus, the closure phase makes it possible to build a rich description of astrophysical sources. However, the Fourier coverage remains sparse. The best solution to this sparsity problem is to use a parametric model of the observed object. To avoid sparsity altogether, one would need to use a denser disposition, such as a regular grid of sub apertures. Such a geometry introduces redundancy, and it is not possible to build closure phases from it in the general case. For small optical default however, there exists a linear relation between small OPDs and the noise on the phases of the visibilities they create.

2.1.3 "Closure phases" from a redundant array

The kernel method relies on modelling a continuous aperture as an array of regularly spaced pupils. This modelling approach is akin to considering a discrete version of the Huygens Fresnel principle, where each point of the aperture is considered to be an independent emitter. In the context of the kernel method, the spacing of these virtual sub pupils is related to the quality of the modelling, as a denser model gives a better description.

For most discrete aperture models, each baseline, or spatial frequency is sampled several times, by many different pairs of pupils. This redundancy places us in a regime where the method of closure phases cannot be ported directly, since it relies on a *non redundant* sampling of the Fourier domain. There is however a small OPD regime in which observables akin to closure phases can be constructed from the phases of the complex visibilities.

Let us take a simple array of four distinct pupils disposed in a square, shown in Fig. 2.3, where 4 distinct baselines are realised by 4 pupils. The baselines $\boldsymbol{a}$ and $\boldsymbol{b}$ are realised two times in this case. If a baseline is realised by several pairs of pupils, the resulting interference pattern on the focal plane is the *incoherent* sum of the interference patterns produced by each pair. Since the Fourier transform is a linear transformation [2], the complex visibility for the corresponding baseline is the sum of the complex visibilities that would be obtained by each of the pairs. For a given baseline $\boldsymbol{l}$, the complex visibility at the corresponding spatial frequency is thus

$$\mathcal{V}(\boldsymbol{l}) = \sum_{k=1}^{N} \mathcal{V}_k(\boldsymbol{l}), \tag{2.13}$$

with the index k corresponding to each of the pupil pairs. This decomposes into the contributions of the object and of the OTF written in Eq. (1.4)

$$\mathcal{V}(\boldsymbol{l}) = \sum_{k=1}^{N} OTF_k(\boldsymbol{l})\mathcal{V}_k^0(\boldsymbol{l}). \tag{2.14}$$

[2] i.e. $\mathcal{F}\big(f(x) + g(x)\big) = \mathcal{F}\big(f(x)\big) + \mathcal{F}\big(g(x)\big)$

Since the visibility of the object is the same for every realisation of the same baseline,

$$\mathcal{V}(\boldsymbol{l}) = \mathcal{V}_0(\boldsymbol{l}) \sum_{k=1}^{N} OTF_k(\boldsymbol{l}). \tag{2.15}$$

As seen in Eq. (2.7), the PTF for each pair of pupils is the phase shift imposed by each pupil. The MTF is the autocorrelation of the pupils, but can be arbitrarily set to one if all the pupils are identical. Noting $\delta\varphi_k$ the phase shifts for each pair of pupils realising the baseline $\boldsymbol{l}$, and using the decomposition of the OTF in Eq. (2.5), we then get

$$\mathcal{V}(\boldsymbol{l}) = \mathcal{V}_0(\boldsymbol{l}) \sum_{k=1}^{N} e^{i\delta\varphi_k}. \tag{2.16}$$

The phase $\phi(\boldsymbol{l})$ of the measured complex visibility is affected by an additive phase noise

$$\phi(\boldsymbol{l}) = \phi_0(\boldsymbol{l}) + \angle \sum_{k=1}^{N} e^{i\delta\varphi_k}, \tag{2.17}$$

where the symbol $\angle$ denotes the argument of the complex number it is placed in front of. This additive noise is a non linear combination of the phase shifts in the general case. If the phase shifts are small, it is dominated by the first order of its polynomial expansion, and approximation

$$\angle \sum_{k=1}^{N} e^{i\delta\varphi_k} \approx \frac{1}{N} \sum_{k=1}^{N} \delta\varphi_k \tag{2.18}$$

can be made. The coefficient $1/N$ reflects the redundancy of the baseline. For $N = 1$, the relation is exact.

Thanks to this approximation, observables akin to closure phases can be constructed if the phase shifts are small enough, i.e. if the OPDs across the pupils are small compared to the observation wavelength. This regime is reached with observations made using high performance AO or space telescopes.

Let us consider the simple 4 pupils configuration represented in Fig. 2.3. The phases of the complex visibilities measured by the baselines $\boldsymbol{a}$ and $\boldsymbol{b}$, $\boldsymbol{c}$ and $\boldsymbol{d}$ are, taking $\varphi_1 := 0$:

$$\begin{aligned}
\phi(\boldsymbol{a}) &= \phi_0(\boldsymbol{a}) + \angle(e^{i(-\varphi_2)} + e^{i(\varphi_4-\varphi_3)}) \\
\phi(\boldsymbol{b}) &\approx \phi_0(\boldsymbol{b}) + \angle(e^{i(-\varphi_4)} + e^{i(\varphi_2-\varphi_3)}) \\
\phi(\boldsymbol{c}) &= \phi_0(\boldsymbol{c}) + + \angle e^{i(\varphi_4+\varphi_3)} \\
\phi(\boldsymbol{d}) &= \phi_0(\boldsymbol{d}) + \angle e^{i(-\varphi_3)}.
\end{aligned} \tag{2.19}$$

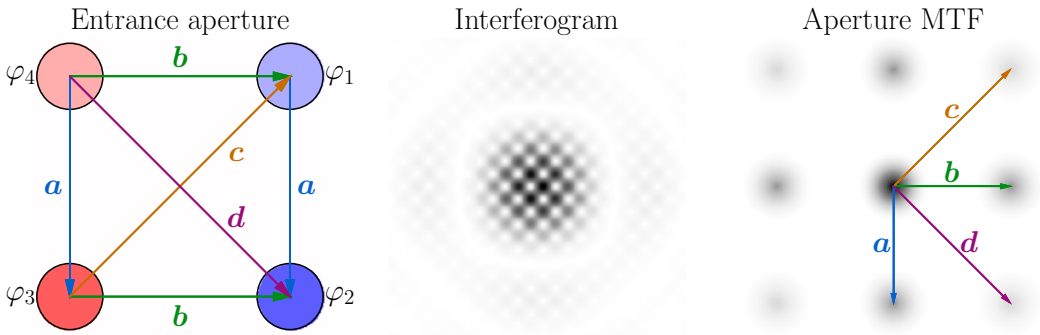

FIGURE 2.3: A redundant, 4 holes interferometer that realises 4 independent baselines. In this case, the baselines $\boldsymbol{a}$ and $\boldsymbol{b}$ are sampled twice. Thus, the complex visibilities sampled at the corresponding spatial frequencies are the sum of the complex visibilities from each pair.

If the phase shifts φ_2, φ_3 and φ_4 are small, the approximation Eq. (2.18) can be used, and the observed phases can be approximated as

$$\begin{aligned}
\phi(\boldsymbol{a}) &\approx \phi_0(\boldsymbol{a}) + \frac{1}{2}(-\varphi_2 + \varphi_4 - \varphi_3) \\
\phi(\boldsymbol{b}) &\approx \phi_0(\boldsymbol{b}) + \frac{1}{2}(-\varphi_4 + \varphi_2 - \varphi_3) \\
\phi(\boldsymbol{c}) &= \phi_0(\boldsymbol{c}) + \varphi_4 - \varphi_2 \\
\phi(\boldsymbol{d}) &= \phi_0(\boldsymbol{d}) - \varphi_3.
\end{aligned} \tag{2.20}$$

Using this now linear relation, it is possible to find combinations of the phases of the complex visibilities for which the aberration terms cancel each other out. The combinations

$$\begin{aligned}
k_1 &= 2\phi(\boldsymbol{a}) + 2\phi(\boldsymbol{b}) + \phi(\boldsymbol{d}) \\
k_2 &= 2\phi(\boldsymbol{b}) + \phi(\boldsymbol{c}) - 2\phi(\boldsymbol{d}) \\
k_3 &= 2\phi(\boldsymbol{a}) - 2\phi(\boldsymbol{b}) - \phi(\boldsymbol{c})
\end{aligned} \tag{2.21}$$

are independent from the error terms φ_2, φ_3 and φ_4 at the first order. These observables, which depend on the linear approximation of the phase of the sum of complex numbers are called kernel phases.

In this example, the relation that yields a kernel phase can be arrived at "by hand", but in practice a redundant array is a pupil *model*, that comprises hundreds or even thousands of pupils and of redundant baselines. To derive kernel phase using such a model, a more systematic approach is needed.

2.2 Kernel phases

2.2.1 Construction of kernel phases from an arbitrary array using linear algebra

The construction of kernel phases relies on the determination of the linear relations that approximate the propagation of the OPD (that causes phase shifts across the apertures), to the phase noise on the visibility. Rather than write a full system of equations, linear algebra can be employed to condense these relations into a single one.

Looking at the example in the previous section, the relation Eq. (2.20) can be rewritten by defining the measured phases vector

$$\boldsymbol{\phi} = \begin{bmatrix} \phi(\boldsymbol{a}) \\ \phi(\boldsymbol{b}) \\ \phi(\boldsymbol{c}) \\ \phi(\boldsymbol{d}) \end{bmatrix} \tag{2.22}$$

while the phase shifts on each pupil are collected in the vector

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \end{bmatrix}. \tag{2.23}$$

The linear approximations linking the phase shifts on each pupil to the phases of the complex visibilities Eq. (2.20) can then be written

$$\boldsymbol{\phi} \approx \boldsymbol{\phi}_0 + \boldsymbol{A}\boldsymbol{\varphi}, \tag{2.24}$$

with $\boldsymbol{A}$ the *transfer matrix*

$$\boldsymbol{A} = \begin{bmatrix} -\frac{1}{2} & +\frac{1}{2} & -\frac{1}{2} \\ +\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{2.25}$$

A set of kernel phases for this particular configuration can be found using an algebraic kernel of the transfer matrix $\boldsymbol{A}$, which us a base of the null space of this matrix. Martinache (2010) uses the singular value decomposition of $\boldsymbol{A}$,

$$\boldsymbol{A}^T = \boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^T. \tag{2.26}$$

The kernels matrix $\boldsymbol{K}$ is constructed by taking the columns of $\boldsymbol{V}^T$ that correspond to singular values (the diagonal terms of $\boldsymbol{W}$) equal to zero, and for an

orthogonal base of $\boldsymbol{A}$.[3].

For the purpose of creating a set of observables independent from OPDs across the aperture, the important property of $\boldsymbol{K}$ is that it is the largest matrix with linear independent rows such as

$$\boldsymbol{KA} = \boldsymbol{0}, \tag{2.27}$$

and therefore, the set of kernel phases defined as

$$\boldsymbol{k} := \boldsymbol{K\phi} \tag{2.28}$$

is independent from the first order effects of the OPDs across the aperture, giving, in the absence of noise in the image,

$$\boldsymbol{k} \approx \boldsymbol{K\phi}_0. \tag{2.29}$$

Kernel phases can be easily constructed from an image with the appropriate pixel sampling (more than two pixels per λ/D element) and the geometry of the entrance aperture using the `xara` Python package[4], which handles the construction of the aperture model, of the transfer matrix and of the kernels matrix. In the past years, the modelling of the aperture has been improved to include pupils with different transmissions to better describe the edges and fine details of the aperture (Martinache et al., 2020).

Since its emergence ten years ago, the kernel method has been used on images from multiple large ground based telescopes, as well as from the Hubble Space Telescope (Pope, Martinache, and Tuthill, 2013; Laugier et al., 2019a; Kammerer et al., 2019; Pope, 2016). These practical uses have revealed the presence of multiple stellar and substellar binaries, and allowed a precise measurement of their parameters. So far, the method has focused on the detection of stellar and substellar companions: a binary can be modelled with three parameters, making the inverse problem less difficult to solve than for more complex structures, such as protoplanetary disks. These objects are also of scientific interest: the multiplicity of brown dwarfs for instance is not well know (Leggett et al., 2017; Fontanive et al., 2018). Furthermore, measurement of the relative positions of two objects enables the reconstruction of their orbits and therefore the determination of their dynamical masses. This is of particular interest for the determination of mass luminosity relations, to constrain evolution models (Burrows et al., 2001). So far, the method has focused on the detection of companions at contrasts typically below 10^2. More ambitious programs push toward detection at higher contrast.

[3]Martinache (2010) separates the transfer matrix into a diagonal redundancy matrix and a transfer matrix filled with values that are -1, 0 or 1. This approach has numerical advantages for the computation of $\boldsymbol{K}$, but the simplified approach presented here leads to a kernels matrix $\boldsymbol{K}$ that lies in the same subspace, and is therefore equivalent.

[4]available at https://github.com/fmartinache/xara.

2.2.2 Kernel amplitudes

A linear model of the propagation of the OPDs from an array of pupils to the *amplitudes* of the complex visibilities can be constructed using the dispositions of the pupils, as shown by Pope (2016). This approach enables the recovery of all of the complex visibility information that can be separated from the OPDs across a telescope's entrance aperture. This approach relies on the transformation of the multiplicative relation between the complex visibility and the OTF into a sum, by taking the logarithm of the measured complex visibility.

This approach does work in practice, and produces kernel amplitudes largely independent from the phase shifts across the entrance aperture. However, the use of a logarithm means the relation between the pixel values across the images and the kernel amplitudes is largely non linear, which complicates the statistical analysis of these observables. This dissertation focuses on the treatment of kernel phases, but adding amplitudes to the observables is an exciting avenue for improving the method and obtain richer observables, from which it may in principle be possible to reconstruct images that have been "cleaned“ from the contributions of OPDs across the aperture.

2.3 Kernel phases errors

Since kernel and closure phases are constructed to be robust to optical defaults, one would expect the kernel phases extracted from the image of a single star with a very small angular diameter to be zero. In practice however, this is not what is observed, and kernel phases are never perfectly null. The origin of this residual signal can be broken down into two categories. Firstly, pixel level noise in the image, which also affects its Fourier transform, and thus the complex visibilities extracted from said image. The other sources of residuals are not as easy to pinpoint, as they result from a combination of many factors, from atmospheric dispersion to aberration residuals.

Here, I provide an overview of these error sources, drawing from previous work by Ireland (2013), and how they can be estimated, and ultimately mitigated.

2.3.1 Systematic errors

Although the kernel transformation does greatly reduce the contribution of the OPDs in the kernel phases observables, it does not completely eliminate them. Some optical defaults in particular can be considered systematic, and consist of effects such as dispersion, low order residual aberrations in the optical systems, or stray light due to unwanted reflections that create a weak parasitic signal in the kernel phases. These effects make up systematic errors. Some on these effects remain constant throughout the observation, and thus the biases they introduce can be corrected via calibration: subtracting an

estimation of the bias from the measured kernel phases allows it to be greatly reduced, or even virtually eliminated if the defaults in the optical path do not change too greatly between the acquisition of the calibration frames and of the on-target frames. A typical estimation method is to use the kernel phases obtained from an unresolved source (Pope et al., 2016).

Calibration does not necessarily provide a perfect correction of systematic errors. Calibration frames also comprise pixel level noises, adding to the noise affecting the science frames, and the biases are not necessarily stable through time. For instance, aberrations evolve, as thermal effects lead to deformation of the structure of the telescope, and the elevation of the target may change over time, leading to a variation in atmospheric dispersion. A possible solution to this issue is to commute as quickly as possible between the object of interest and a calibration target, offering very frequently updated estimates for the bias, at timescales smaller than those over which it evolves. This is the approach of chopping employed in the NEAR instrument, were the image commutes rapidly between a calibrator and a scientific source, such as Käufl et al. (2018) for α Centauri. However, this technique is only usable in a narrow range of observing conditions, as it requires the scientific target and the calibrator to be extremely close from each other.

The field rotation during an observation can also be used to obtain self calibrating kernel observables (Laugier et al., 2019a). In this case, it is knowing how the object of interest rotates in the sky, while the bias remains unaffected by this rotation that enables the measurement of the bias signal, without the need for a calibrator. This angular differential methods in inspired by angular differential imaging (Marois et al., 2006), which employs field rotation to remove coronagraphic residuals.

Other methods can be employed to estimate the bias, with linear combinations of different calibration signals, as suggested by Ireland (2013), and recently applied to observations by Kammerer et al. (2019). There, the bias is considered to be a linear combination of several components, and estimating it is equivalent to estimating the relative importance of each of those components.

In practice, kernel biases are always present. If they are small enough compared to the contribution of the noise, they can be neglected safely. However, if their contribution is more important, they must be taken into account statistically. This entails an estimation of the distribution of the unknown biases. This no easy task: since the phenomena that contribute to them evolve simultaneously on a variety of timescales, making any estimation of their distribution at a given time difficult. This is especially true for conditions that can vary greatly during an observing run. For instance, atmospheric turbulence can evolve a lot during a single night of observation, and the distribution of the bias it causes at one time it not necessarily representative of what it can become a few hours, or even minutes later. This problem, compounded by the number of phenomena that can lead to biases make the estimation of their distributions very hard.

Residual optical defaults are mostly corrected thanks to calibration. Since these residuals are closely tied to the pupil model, alterations of this discrete model to make it more representative of the discrete reality can help in reducing their contribution. The way this has been done is by taking the partial obstruction of some of the subapertures that make up the model by structures in the apertures into account. as presented in Martinache et al. (2020).

2.3.2 Noisy observables from noisy images: statistical errors.

The errors that result from the noise in the image cannot be reduced by improving the extraction of kernel phases, contrary to systematic errors. On the flip side, they are easier to estimate, as the noise produced for one given exposure is largely independent from the noise produced in any other exposure, assuming the aberration remains the same over the exposure.

The noise in the image is caused both by the natural fluctuation in the number of photons that hit each pixel in a given time, and by the accuracy with which the detector can measure this quantity. The distributions of these errors are determined by the inherent characteristics of the detector, as well as by the flux received by each pixel. Pixel level errors can be broken down into the following categories:

- Photon noise: Poisson distributed, and caused by the discrete nature of light. With many photons, the Poisson distribution becomes similar to a normal distribution. (Station, 1957).
- Read noise: Normally distributed, and caused by errors in the reading process of each pixel's voltage (Mccord and Bosel, 1975).
- Dark current noise: photodetectors accumulate electrical charge, even in the absence of light. This creates a bias, that can be corrected by subtracting a dark frame, but also introduces an additional Poisson distributed dark current noise, since the electrons accumulate in each well following a Poisson process.
- Quantization noise: Most detectors used in astronomy employ 16 bits registers, meaning each pixel value for the detector can take only $2^{16} = 65536$ discrete values.

Although noise is pixel independent (excluding bad pixels, for which the measured flux is interpolated from the measurements at the level of nearby pixels), the measured complex visibilities, and thus kernel phases are not.

At the level of each pixel, the measured flux follows a mix of Poisson and normal distributions (the distribution function of the sum two random variables is the convolution of the distribution functions of each). However, kernel phases are the result of the combination of the values of all of the pixels in the image. Complex visibilities are obtained via a matrix multiplication that performs the discrete Fourier transform of the image (see Martinache

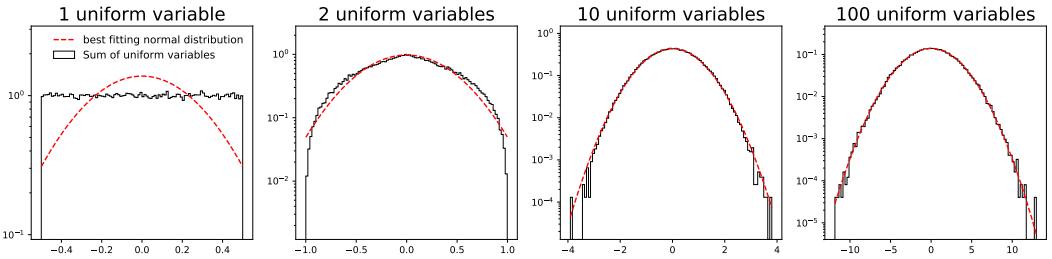

FIGURE 2.4: A linear combination of many random variables with distributions of finite variance and mean tends to follow a normal distribution.

(2018), pp 197-199). The arctangent of the ratio between the imaginary part and the real part of the visibility then serves to measure its phase:

$$\angle(a + ib) = \arctan \frac{b}{a}. \tag{2.30}$$

For small fluctuations of the phase, this operation can be approximated as linear. Then, the measured phases vector is multiplied by the kernels matrix: each kernel phases is a linear combination of the complex phases measured for each spatial frequency sampled by the discrete pupil model.

The central limit theorem states that a linear combination of a high number of independent random variables with distributions of finite variances and means tends towards being normally distributed, as illustrated in Fig. 2.4. Since kernel phases are the result of a linear combination of random variables, and since a pupil model typically samples a large number of spatial frequencies, one can therefore expect kernel phases to be normally distributed. If each kernel phase is normally distributed, then the kernel phases vector can be described as following a multivariate, normal distribution:

$$\boldsymbol{k} \sim \mathcal{N}(\boldsymbol{k}_0, \boldsymbol{\Sigma}), \tag{2.31}$$

with $\boldsymbol{k}_0$ the expected kernel phase value (including the object signature and any possible residual bias), and $\boldsymbol{\Sigma}$ the covariance matrix of the noise propagated to the kernel phases, which encodes both the variance and the covariance of the kernel phases. Since the noise is normal, the distribution of the noise is fully described by this covariance matrix.

Knowing the distribution of the kernel phases (or of any other measured quantity for that matter) is important when it comes to their statistical treatment: this knowledge is fundamental to determining how likely a given object signature is for a particular set of measured kernel phases. It is also important to determine the likelihood one would obtain a given set of kernel phases from random fluctuations in the noise, rather than from an object signal. This latter aspect is crucial to the reliability of detection. To detect a signature, it is necessary to know how likely one can obtain a similar signal

from random fluctuations of the noise.

The detection procedures outlined in the next chapter of this thesis rely on the distribution of the kernel phases to construct detection tests that can be applied to kernel phases, and help detect either arbitrary signatures, unspecified signatures, or binaries. To be able to perform detections on any sets of kernel phases, the noise is whitened, i.e.kernel phases are normalised using the covariance of the nose, allowing all possible kernel phases to have the same distribution.

From images taken with any telescope that can reach a regime where aberrations are low enough, it is possible to construct observables that are largely unaffected by the residual aberrations. This enables the detection of faint features around star or substellar objects at very close separations. Attempting detections at more and more ambitious separations and contrasts does however mean the statistical treatment of kernel observables must take errors into account with care, as the chance of making a detection wrongfully increases when the signals one attempts to measure become weaker, such as for companions at close separations and high contrasts. By estimating both the distribution of each kernel phase, and the correlations that exist between them, it is possible to devise statistical detection tests, which can offer both guarantees of optimality (being the most powerful detection test that can be devised for a given signal and noise level), as well as guarantees over the likelihood of a companion or structure being present.

2.4 Conclusion

In the age of space telescopes and extreme adaptive optics, the minimisation of the contribution of optical defaults via post acquisition treatment remain important to probe small angular separations, as well as faint structures. For instance, high contrast imaging combines advanced coronagraphs and post treatment to minimise the impact of speckles on the final image (e.g Cantalloube et al. (2015)).

Kernel phases also permit the use of full aperture images, but without using a coronagraph, which enables the probing of regions very close to a star. Their extraction does not require a non redundant mask, which improves the available flux, and thus reduces the contribution of photon noise.

Knowing the distributions of kernel phases enables the determination of the likelihood of an object having a given signature from an image. This is the key to making detections and estimating the signature of an observed object. The statistical treatment of kernel phases is the core of the work I present in this manuscript, and the focus of the next chapter.

Chapter 3

Statistical detection of binaries

To be able to draw conclusions from a measurement affected by random errors, one must use statistical methods. In the case that concerns us here, I focus on the detection of companions from the kernel phases extracted from images.

In this chapter, I first present, in Sec. 3.1 an introduction to the field of statistical detection, and the essential notions and terminology that will be employed throughout Sec. 3.2. In this section, I show how systematic, likelihood ratios approaches can be employed to design statistical tests, first explaining how the distribution of the test statistic is linked to its performances in Sec. 3.2.1, and then going on to formalise three tests. The first one, introduced in Sec. 3.2.2, is tuned to detect a specific signal (an object with an already known structure). It is the most specific possible test, which is also the performance benchmark, since it has the highest possible power as demonstrated in Appendix A. Sec. 3.2.3 and Sec. 3.2.4 presents two other tests, that aim to detect signatures of an unspecified structure for the former, and of a binary for the latter. Sec. 3.2.5 presents a brief discussion of the relation between these tests and "goodness of fit" evaluations used on normally distributed data.

The following part of this chapter is dedicated to a practical use case for these detection tests for the statistical treatment of kernel phases obtained with the NIRISS instrument on JWST in Sec. 3.3. I present the scientific use case of Y-type brown dwarfs in Sec. 3.3.1, the errors one can expect on kernel phases extracted from JWST NIRISS images in Sec. 3.3.2, the way parameters of binaries can be measured from these kernel phases in Sec. 3.3.4, with the Appendix B proposing an alternative method to obtain these parameters.

The following Section, Sec. 3.4 dives into the predicted performances of the aforementioned test for the purpose of detection companions of faint Y type brown dwarfs using JWST NIRISS images. Sec. 3.4.1 directly compares the three tests' power, Sec. 3.4.2 translates those detection limits into mass and absolute separation limits for a known Y dwarf, and Sec. 3.4.3 gives the ultimate possible contrast limits for a kernel phases based detection using JWST NIRISS images with the F480M filter.

In the conclusion, I remark how detection can be made well below the λ/D "diffraction limit", and follow up on this observation in Appendix C, where

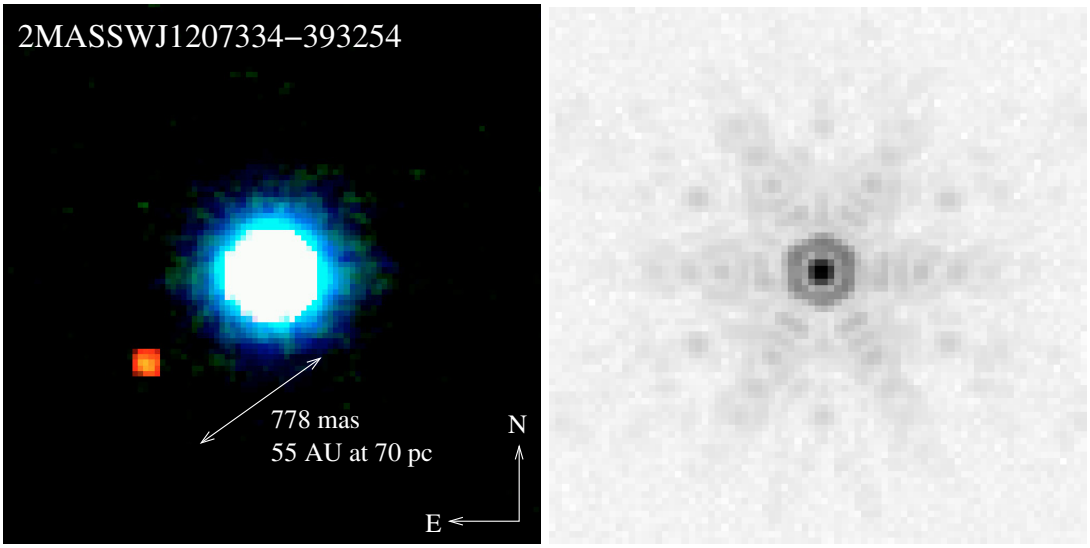

FIGURE 3.1: Left: VLT/NACO image of the 2M1207 system (taken from Chauvin et al. (2004)). Right: Synthetic JWST NIRISS image, with a companion injected at a separation of 147 mas. Detecting 2M1207b does not require advanced signal processing, as its presence is very obvious in the image. In more challenging cases however, a visual confirmation is not sufficient, although, as we will show at the end of this chapter, in Sec. 3.5, detecting a companion at a separation of 147 mas and a contrast of 50 becomes easy using kernel interferometry and the proper statistical tools.

I show how the phase signature of a binary becoming degenerate with a tip tilt aberration limits the separations that an be probed using kernel phases.

The theoretical framework, as well as the predicted detection limits for JWST NIRISS images presented in this chapter have been the object of a publication in Astronomy & Astrophysics: Ceau et al. (2019), reproduced at the end of this manuscript.

3.1 Statistical detection

The less obvious the presence of a signal, the more crucial the statistical treatment of the data to reveal it must be. In the case of images, some features can for instance be detected by a visual inspection, but kernel phases can be used to detect features that are completely invisible when looking at images. As an example, the simulated JWST NIRISS image on the right hand side of Fig. 3.1 features a companion in the first diffraction ring, at a contrast of 50, which can be detected with ease using the tests I present in Sec. 3.2.

Statistical detection aims at distinguishing between hypotheses (typically, "signal present" and "signal absent") in noisy data. It is a fundamental problem of signal processing, with many ramifications, ranging from military applications (e.g. radar signatures Marcum (1947)), healthcare (e.g. the early diagnosis of conditions such as cancer or HIV infection), and many other topics. Here, we most specifically use binary hypothesis tests, where data serves to decide between two competing hypotheses: a null hypothesis, noted $\mathcal{H}_0$ and an alternative hypothesis, noted $\mathcal{H}_1$.

For a given measurement, a hypothesis test returns a test statistic, which is a real number here. The test statistic is compared to a threshold, which sets the significance level. Knowing the distribution of the test statistic under $\mathcal{H}_0$ is necessary to determine the threshold, needed to obtain a targeted false alarm rate. The false alarm rate is the probability that the test returns a detection (i.e, that it chooses $\mathcal{H}_1$ while $\mathcal{H}_0$ is true).

In this chapter, I focus on the conception and use of binary hypothesis tests for kernel-phases extracted from images. More specifically, I show how the kernel phases obtained from the image of a single source (which are expected to be distributed around zero, as discussed in Sec. 2.1.2), such as a star can be distinguished from the kernel phases originating from a more complex object (a binary in this case). This exploits the fact that the phase of the complex visibility of any centrosymmetric object is zero, while asymmetrical features break this symmetry, and create a phase signal in the complex visibility, as explained in Sec. 2.1.2. Kernel phases, being largely immune to the phase errors caused by aberrations are thus distributed around zero if the imaged object is a single source, and around non zero kernel phases if the object has some asymmetrical features.

The statistical tests that enable the detection of such features involve a mathematical transformation which turns the data into a test statistic. Knowing the distribution of the test statistic under $\mathcal{H}_0$ necessitates knowing the distribution of the noise that affects the data. In the case of kernel phases, it means determining the distribution of the different noises that affect these observables. As seen in Chapter 2, Kernel phases can be assumed to be correlated with each other, and normally distributed. Their distribution can be described using a multivariate, normal distribution, which is characterised by its mean and its covariance matrix.

The detection procedures used here are based on likelihood ratios. To determine which of two hypotheses must be decided, the likelihoods of obtaining the observed data given each hypothesis are compared using the ratio of the two, and it is this ratio that provides the test statistic. We propose three tests, each corresponding to a given degree of knowledge about the object under $\mathcal{H}_1$. Two of these tests are specific to the detection of stellar or substellar companions (they correspond to the cases where the position and contrast of a companion are known, or unknown). The other test makes no assumption on the target.

3.2 Hypothesis tests

Hypothesis testing is the framework from which we build our detection procedures. From the kernel phases vector $\boldsymbol{k}$, the whitened kernel phases vector $\boldsymbol{y}$ is obtained using the whitening transformation

$$\boldsymbol{y} := \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{k}, \tag{3.1}$$

given the normal distribution of the kernel phases (2.31). We consider the null hypothesis to be the absence of asymmetrical features, which create no kernel phase signal (in which case the data is only comprised of noise), and the alternative hypothesis the presence of such features which creates a signature $\boldsymbol{k}_0 = \boldsymbol{K}\boldsymbol{\phi}_0$, as seen in Eq. (2.28) in the kernel phases, whitened into

$$\boldsymbol{x} := \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{k}_0. \tag{3.2}$$

The two cases for the whitened measure kernel phases are then noted:

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y} = \boldsymbol{\epsilon} \\ \mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon} \end{cases} \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}) \tag{3.3}$$

with $\boldsymbol{\epsilon}$ the noise and $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ the standard, multivariate normal distribution. Using whitened observables allows us to use the generic description in Eq. (3.3) for the noise for any possible covariance, as long as the noise that affects the observables vector is normally distributed and its covariance is known.

A *detection* is a rejection of the null hypothesis. In the case of kernel phases analysis, this corresponds to the detection of the whitened kernel phase signature $\boldsymbol{x}$ from the data $\boldsymbol{y}$.

As outlined above, a detection test is a function that, from some input data returns a test statistic, i.e. a single, real number. The false alarm rate is given by the distribution of the test statistic under the null hypothesis. To control the false alarm rate, the distribution of the test statistic under the null hypothesis therefore needs to be known. The decision between $\mathcal{H}_0$ and $\mathcal{H}_1$ is then made by comparing the test statistic $T(\boldsymbol{y})$ to a threshold ξ, which is the minimum value the test statistic must take to reject the null hypothesis.

$$T(\boldsymbol{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi \tag{3.4}$$

The test thus chooses $\mathcal{H}_0$ if the test statistic is smaller than ξ and $\mathcal{H}_1$ otherwise. The false alarm rate P_{FA} and the detection rate P_{DET} are linked to the test statistic and to the thresholds by the relations

$$\begin{aligned} P_{FA}(\xi) &:= P(T(\boldsymbol{y}) > \xi; \mathcal{H}_0) \\ P_{DET}(\xi) &:= P(T(\boldsymbol{y}) > \xi; \mathcal{H}_1) \end{aligned} \tag{3.5}$$

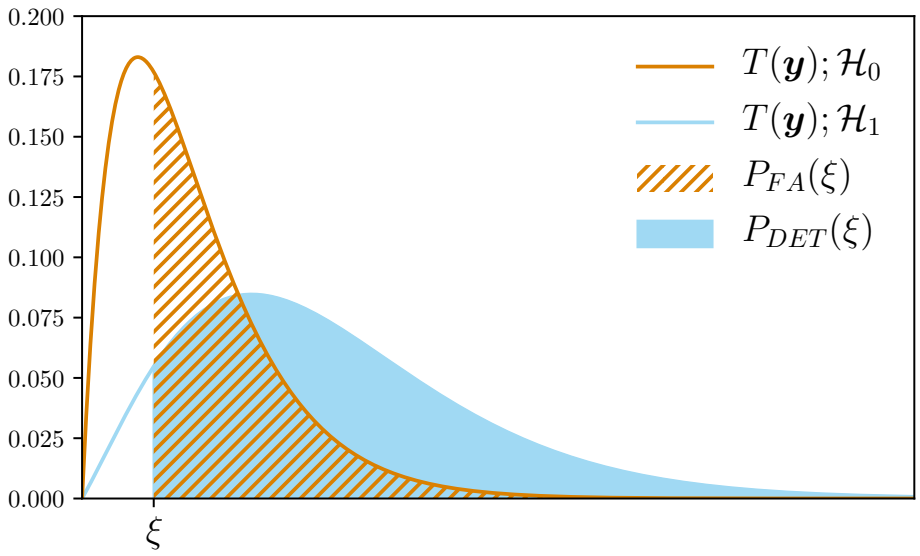

FIGURE 3.2: Illustration of the detection and false alarm rates for a given test. The distributions of the test statistic under $\mathcal{H}_0$ and $\mathcal{H}_1$ are fixed. The two solid lines represent the distributions of the test statistic under $\mathcal{H}_0$ and $\mathcal{H}_1$, while the filled areas represent P_{FA} and P_{DET}.

The false alarm rate and detection rate are graphically illustrated in Fig. 3.2. In practice, detections are made at a given significance level, which is related to the false alarm rate: the lower the false alarm rate, the higher the significance level.

The power of a test is given by its detection rate, or P_{DET} at a given P_{FA}: the higher the probability to make a detection at a given signal and false alarm rate, the more powerful the test. The power of a test can be captured by its ROC, for receiver operating characteristic, which is the function that links P_{FA} to P_{DET} for a given test.

In the illustration in Fig. 3.2, this means the more separated the two distributions of the test statistic under $\mathcal{H}_0$ and $\mathcal{H}_1$, the more powerful the test. Looking at the distributions on a plot is not the most effective way to compare the powers of tests. The distribution in Fig. 3.2 can be condensed as a relation between P_{FA} and P_{DET} as the threshold varies in the form of the receiver operating characteristic, or ROC: the more powerful a test, the higher its P_{DET} for a given P_{FA}. ROC curves allow to compare tests at a glance, and to condense the power information to a single curve per test. An example of ROC curves for three tests with different powers is shown in Fig. 3.3.

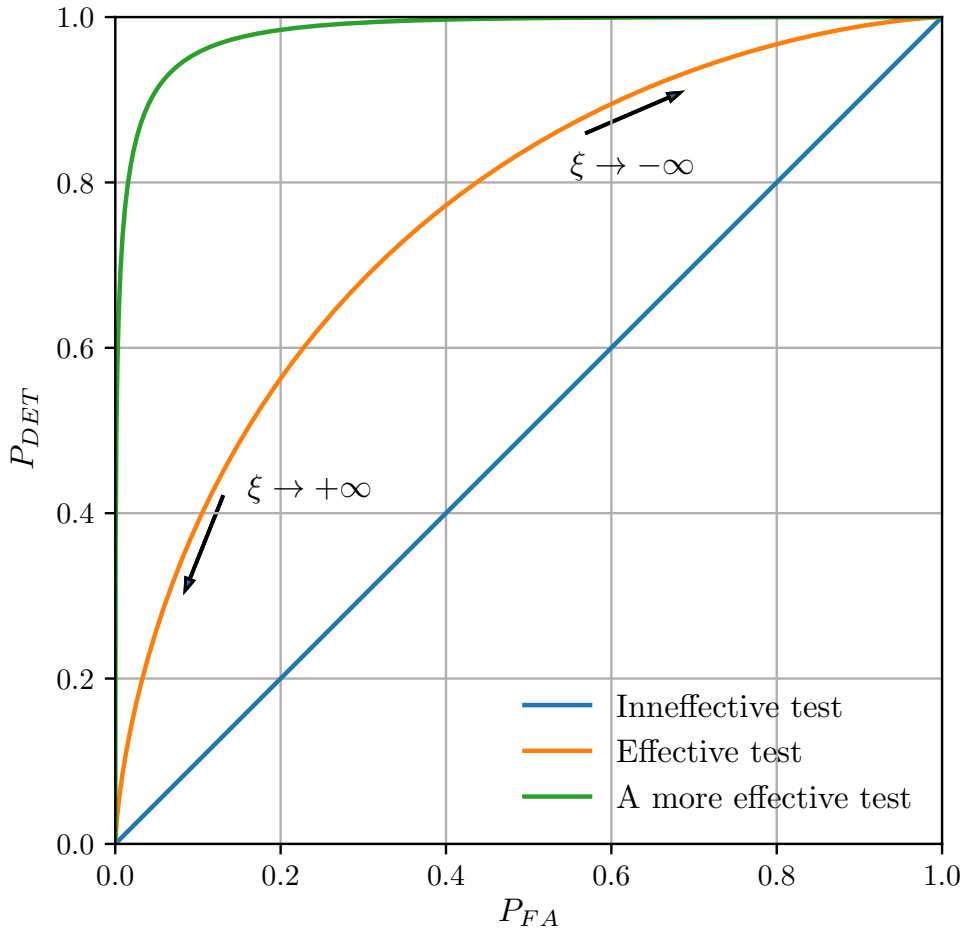

FIGURE 3.3: A ROC curve can be used to compare the performances of several tests. It consists in plotting, for a given test and signature the P_{FA} and the P_{DET} as the threshold ξ changes. The power of a test is its P_{DET} for a given P_{FA}. Visually, a higher power translates as the ROC curve being "pulled" into the upper left corner of the plot. If a curve is on the diagonal, this indicates that $P_{FA} = P_{DET}$ for every threshold: such a test is equivalent to making decisions at random, and has the worst possible power.

3.2.1 Distribution function and ROC

If the distribution of a test can be analytically described by a distribution function under both $\mathcal{H}_0$ and $\mathcal{H}_1$, then the ROC that describes the performance of the corresponding test also has an analytical form.

The continuous distribution function (CDF) of a random variable R is defined as

$$\mathcal{F}_R(\xi) := P(R \leq \xi), \tag{3.6}$$

For a hypothesis test

$$T(\boldsymbol{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi \tag{3.7}$$

if the distributions functions for the test statistic under $\mathcal{H}_0$ and $\mathcal{H}_1$ are

$$\begin{cases} \mathcal{F}_{T,0} := P\Big(T(\boldsymbol{y}) < \xi; \mathcal{H}_0\Big) \\ \mathcal{F}_{T,1} := P(T\Big(\boldsymbol{y}) < \xi; \mathcal{H}_1\Big) \end{cases} \tag{3.8}$$

then, recalling the definitions of P_{FA} and P_{DET} from Eq. (3.5), the false alarm and detection rates can be rewritten in terms of these distribution functions, with

$$\begin{aligned} P_{FA}^T &= P\Big(T(\boldsymbol{y}) > \xi; \mathcal{H}_0\Big) \\ &= 1 - P\Big(T(\boldsymbol{y}) < \xi; \mathcal{H}_0\Big) \\ &= 1 - \mathcal{F}_{T,0}(\xi) \end{aligned} \tag{3.9}$$

and

$$\begin{aligned} P_{DET}^T &= P\Big(T(\boldsymbol{y}) \geq \xi; \mathcal{H}_1\Big) \\ &= 1 - \mathcal{F}_{T,1}(\xi). \end{aligned} \tag{3.10}$$

The ROC of a test is defined as $P_{DET}(P_{FA})$. The threshold can be expressed as a function of the false alarm rate using Eq. (3.9), giving

$$\xi = \mathcal{F}_{T,0}^{-1}(1 - P_{FA}^T). \tag{3.11}$$

Thus, replacing ξ with this expression in Eq. (3.10) gives

$$P_{DET}^T(P_{FA}) = 1 - \mathcal{F}_{T,1}\Big(\mathcal{F}_{T,0}^{-1}(1 - P_{FA}^T)\Big) \tag{3.12}$$

which is the analytical relation that links the ROC, or the performance of the test to the distribution of the test statistic $T(\boldsymbol{y})$ under $\mathcal{H}_0$ and $\mathcal{H}_1$.

In the context of kernel-based object detection, different tests can be deployed for different sets of hypotheses. I propose detection tests for three different cases: a fully known signature, a fully unknown signature, and a signature of a known structure (a binary system) with the relative positions of the two objects and their contrasts unknown. For the first of theses cases

(known signature), there is a test that has the highest possible power, the likelihood ratio (LR) test. This test gives us the maximum theoretical detection limits for a given signature at any given noise level. The two other tests rely on an estimation of the signature $\boldsymbol{x}$.

3.2.2 Known signature

This case corresponds to the detection of a signature that is known in advance. In this case, there exists an optimal detection test, with the highest power theoretically attainable. It is also the mathematically simplest test to devise, since the signature doesn't have to be estimated from the measured data. Here, we examine a test between the two hypotheses:

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y} = \boldsymbol{\epsilon} \\ \mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon} \end{cases} \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \tag{3.13}$$

with $\boldsymbol{x}$ *known*.

In this situation, the most powerful detection test is a likelihood ratio, or Neyman-Pearson test, where the test statistic is the ratio of the likelihoods of the null and alternate hypotheses given the data under test. The optimality of the likelihood ratio detection test is guaranteed by the Neyman-Pearson lemma (Neyman and Pearson, 1933), a proof of which is given in Appendix A.

To obtain the Neyman-Pearson test statistic, it is necessary to compute the likelihoods $\ell(\boldsymbol{0}; \boldsymbol{y})$ and $\ell(\boldsymbol{x}; \boldsymbol{y})$. Since the noise $\boldsymbol{\epsilon}$ follows a standard normal multivariate distribution with p dimensions, (p corresponds to the number of measured ker-phases), we have the likelihoods:

$$\begin{aligned} \ell(\boldsymbol{0}; \boldsymbol{y}) &= (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{y}\right) \\ \ell(\boldsymbol{x}; \boldsymbol{y}) &= (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^T(\boldsymbol{x}-\boldsymbol{y})\right). \end{aligned} \tag{3.14}$$

The Neyman-Pearson test is:

$$T_{NP}(\boldsymbol{y}, \boldsymbol{x}) := \frac{\ell(\boldsymbol{x}; \boldsymbol{y})}{\ell(\boldsymbol{0}; \boldsymbol{y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{3.15}$$

With the likelihoods in Eq. (3.14), this gives

$$T_{NP}(\boldsymbol{y}, \boldsymbol{x}) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^T(\boldsymbol{x}-\boldsymbol{y})\right)}{\exp\left(-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{y}\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{3.16}$$

A strictly increasing function, (say $f(x)$) applied to the test statistic does not affect the power of the test, as the rejection region is the same for T and $f(T)$.

Since the logarithm is strictly increasing, and defined for every possible value of $(\boldsymbol{x}-\boldsymbol{y})^T(\boldsymbol{x}-\boldsymbol{y})$ and of $\boldsymbol{y}^T\boldsymbol{y}$, it is possible to turn the Neyman-Person test in Eq. (3.16) into the test

$$\begin{aligned} T_{NP}(\boldsymbol{y},\boldsymbol{x}) &= -(\boldsymbol{x}-\boldsymbol{y})^T(\boldsymbol{x}-\boldsymbol{y}) + \boldsymbol{y}^T\boldsymbol{y} \\ &= 2\boldsymbol{x}^T\boldsymbol{y} - \boldsymbol{x}^T\boldsymbol{x}. \end{aligned} \tag{3.17}$$

Taking the threshold $\eta := 2(\ln\xi + \boldsymbol{x}^T\boldsymbol{x})$ since $\boldsymbol{x}$ is known, the test becomes

$$T_{\mathrm{NP}}(\boldsymbol{y},\boldsymbol{x}) = \boldsymbol{y}^T\boldsymbol{x} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{3.18}$$

Since the distribution of $\boldsymbol{y}$ is known, one can derive the distribution of the statistic of $T_{NP}(\boldsymbol{y},\boldsymbol{x})$. This enables us to find the theoretical relation linking P_{FA} and P_{DET}, which in turn enables the analytical determination of the performance of the test.

Under the null hypothesis, the data vector only contains noise, thus,

$$T_{\mathrm{NP}}(\boldsymbol{y},\boldsymbol{x};\mathcal{H}_0) = \boldsymbol{\epsilon}^T\boldsymbol{x}. \tag{3.19}$$

Since $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I})$, $\boldsymbol{x}^T\boldsymbol{\epsilon}$ is normally distributed, with a zero mean, and a variance equal to $\boldsymbol{x}^T\boldsymbol{x}$. The resulting distribution is

$$\boldsymbol{y}^T\boldsymbol{x} = \boldsymbol{\epsilon}^T\boldsymbol{x} \sim \mathcal{N}(0,\boldsymbol{x}^T\boldsymbol{x}). \tag{3.20}$$

Under $\mathcal{H}_1$, the test statistic becomes

$$\begin{aligned} T_{\mathrm{NP}}(\boldsymbol{y},\boldsymbol{x};\mathcal{H}_1) &= (\boldsymbol{x}+\boldsymbol{\epsilon})^T\boldsymbol{x} \\ &= \boldsymbol{\epsilon}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{x}. \end{aligned} \tag{3.21}$$

This test statistic follows a normal distribution of the same variance as under $\mathcal{H}_0$, but with mean $\boldsymbol{x}^T\boldsymbol{x}$. Hence:

$$\begin{cases} \mathcal{H}_0 : T_{\mathrm{NP}}(\boldsymbol{y}) \sim \mathcal{N}\left(0,\boldsymbol{x}^T\boldsymbol{x}\right), \\ \mathcal{H}_1 : T_{\mathrm{NP}}(\boldsymbol{y},\boldsymbol{x}) \sim \mathcal{N}\left(\boldsymbol{x}^T\boldsymbol{x},\boldsymbol{x}^T\boldsymbol{x}\right). \end{cases} \tag{3.22}$$

Noting $\mathcal{F}_{\mathcal{N}}$ the distribution function of the normal, standard distribution, the false alarm and detection rates P_{FA} and P_{DET} for this test can be derived from their definitions in Eq. (3.5)

$$\begin{aligned} P_{FA}^{T_{\mathrm{NP}}}(\eta) &= 1 - \mathcal{F}_{\mathcal{N}}\left(\frac{\eta}{\sqrt{\boldsymbol{x}^T\boldsymbol{x}}}\right), \\ P_{DET}^{T_{\mathrm{NP}}}(\eta) &= 1 - \mathcal{F}_{\mathcal{N}}\left(\frac{\eta - \boldsymbol{x}^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^T\boldsymbol{x}}}\right), \end{aligned} \tag{3.23}$$

which, by expressing the threshold ξ as a function of P_{FA} gives

$$\eta = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}\,{\mathcal{F}_{\mathcal{N}}}^{-1}(1 - P_{FA}), \tag{3.24}$$

and yields the ROC:

$$P_{DET}^{T_{\text{NP}}}(P_{FA}^{T_{\text{NP}}}) = 1 - \mathcal{F}_{\mathcal{N}}\left(\mathcal{F}_{\mathcal{N}}^{-1}(1 - P_{FA}^{T_{NP}}) - \sqrt{\boldsymbol{x}^T\boldsymbol{x}}\right). \tag{3.25}$$

This analytical relation gives the ROC of the most powerful detection test possible. Other tests, that do not benefit from *a priori* knowledge of the signature $\boldsymbol{x}$ are expected to have a lower power, meaning that the values for P_{DET} at a given P_{FA} will be lower than for the likelihood ratio.

3.2.3 Completely unknown signature

After examining the case where the signature one is searching for is perfectly known (which provides the upper performance bound thanks to the Neyman-Pearson lemma), we now focus on the case where the object's kernel phase signature is perfectly unknown, meaning any element of the vector $\boldsymbol{x}$ can take any value. This case provides the lower bound for power of detection. This can be useful in practice, to detect unspecified features around luminous objects.

For this case, using a likelihood ratio requires the estimation of the object signature, $\boldsymbol{x}$. This can be done by finding the signature which is the maximum likelihood estimate (MLE) $\hat{\boldsymbol{x}}$. The determination of the MLE is an optimisation problem which consists in determining the signature that maximises $\ell(\boldsymbol{x}; \boldsymbol{y})$.

The *generalised likelihood ratio test*, or GLR (Solomon, 1975) is a likelihood ratio, where the unknown likelihood $\ell(\boldsymbol{x}, \boldsymbol{y})$ is replaced by the likelihood $\ell(\hat{\boldsymbol{x}}; \boldsymbol{y})$, where $\hat{\boldsymbol{x}}$ is the MLE

$$\hat{\boldsymbol{x}} := \underset{\boldsymbol{x}}{\text{argmax}}\ \ell(\boldsymbol{x}; \boldsymbol{y}). \tag{3.26}$$

Contrary to the likelihood ratio, or Neyman-Pearson test, the GLR does not carry any proven guarantee of optimality in the general case. However, in some specific cases (Scharf and Friedlander, 1994). .

In the case of an observable affected by an additive, normal multivariate noise of covariance $\boldsymbol{I}$, the likelihood is:

$$\ell(\boldsymbol{x}; \boldsymbol{y}) = (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y})\right). \tag{3.27}$$

The likelihood $\ell(\boldsymbol{x}; \boldsymbol{y})$ is maximal when $(\boldsymbol{x} - \boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y})$ is minimal. Since there is no constraint on the value $\boldsymbol{x}$ is allowed to take (the signature is assumed *completely* unknown), the quantity in the exponential can reach a value of zero for $\hat{\boldsymbol{x}} = \boldsymbol{y}$, which is the MLE in this case.

This MLE is then injected into the likelihood ratio in Eq. (3.15), with $\hat{\boldsymbol{x}}$ substituting $\boldsymbol{x}$:

$$T(\boldsymbol{y}) = \frac{\ell(\hat{\boldsymbol{x}}; \boldsymbol{y})}{\ell(\boldsymbol{0}; \boldsymbol{y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{3.28}$$

The maximum likelihood is

$$\ell(\hat{\boldsymbol{x}}; \boldsymbol{y}) = \ell(\boldsymbol{y}; \boldsymbol{y}) = (2\pi)^{-\frac{p}{2}}, \tag{3.29}$$

and the likelihood of the null hypothesis given the data is

$$\ell(\boldsymbol{0}; \boldsymbol{y}) = (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{y}^T \boldsymbol{y})\right). \tag{3.30}$$

These likelihoods, injected into Eq. (3.28) give the test

$$\begin{aligned} T(\boldsymbol{y}) &= \frac{1}{\exp\left(-\frac{1}{2}(\boldsymbol{y}^T \boldsymbol{y})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi \\ &= \exp\left(\frac{1}{2}(\boldsymbol{y}^T \boldsymbol{y})\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi \\ &= \boldsymbol{y}^T \boldsymbol{y} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} 2 \ln \xi. \end{aligned} \tag{3.31}$$

defining the new threshold $\eta := 2 \ln \xi$. The test then becomes

$$T_{\mathrm{E}}(\boldsymbol{y}, \boldsymbol{x}) = \boldsymbol{y}^T \boldsymbol{y} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{3.32}$$

This particular test is an energy detector (hence the notation T_E). Its test statistic is the squared norm, or "energy" of the whitened observable.

Since the distribution of $\boldsymbol{y}$ is known, the distribution of the test statistic of T_E has an easily determined analytical form under both $\mathcal{H}_0$ and under $\mathcal{H}_1$. Under $\mathcal{H}_0$, since $\boldsymbol{y} = \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ the test statistic of T_E is

$$T_E(\boldsymbol{y}, \boldsymbol{0}) = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} \tag{3.33}$$

which follows a χ^2 distribution, with p degrees of freedom (p being the number of kernel phases). Thus,

$$T_E(\boldsymbol{y}, \boldsymbol{0}) \sim \chi^2_p(0). \tag{3.34}$$

Under $\mathcal{H}_1$, the test statistic is equal to

$$T_E(\boldsymbol{y}, \boldsymbol{x}) = (\boldsymbol{x} + \boldsymbol{\epsilon})^T (\boldsymbol{x} + \boldsymbol{\epsilon}), \tag{3.35}$$

and the test statistic therefore follows a *non-central* χ^2 distribution

$$T_E(\boldsymbol{y}, \boldsymbol{x}) \sim \chi^2_p(\boldsymbol{x}^T \boldsymbol{x}). \tag{3.36}$$

Noting $\mathcal{F}_{\chi^2_p(\lambda)}$ the distribution function of a $\chi^2_p(\lambda)$ distribution with a non centrality parameter λ, the false alarm and detection rates associated with T_E, we can follow a similar reasoning as the one used to obtain Eq. (3.22). Using the definition of the distribution function (Eq. (3.6)) and of the relation Eq. (3.12), the theoretical false alarm and detection rate for the energy detector Eq. (3.32) are:

$$\begin{cases} P^{T_{\mathrm{E}}}_{FA}(\eta) = 1 - \mathcal{F}_{\chi^2_p(0)}(\eta) \\ P^{T_{\mathrm{E}}}_{DET}(\eta) = 1 - \mathcal{F}_{\chi^2_p(\boldsymbol{x}^T\boldsymbol{x})}(\eta). \end{cases} \tag{3.37}$$

The first equation transforms into

$$\eta = \mathcal{F}^{-1}_{\chi^2_p(0)}(1 - P^{T_{\mathrm{E}}}_{FA}), \tag{3.38}$$

which, injected into the expression of $P^{T_{\mathrm{E}}}_{DET}(\xi)$ gives the ROC:

$$P^{T_{\mathrm{E}}}_{DET}(P^{T_{\mathrm{E}}}_{FA}) = 1 - \mathcal{F}_{\chi^2_p(\boldsymbol{x}^T\boldsymbol{x})}\left(\mathcal{F}^{-1}_{\chi^2_p(\mathbf{0})}(1 - P^{T_{\mathrm{E}}}_{FA})\right). \tag{3.39}$$

The energy detector can be used on any data set, without any assumption being made on the structure of the researched signature. This makes it very versatile. Applied to kernel phases, it enables the detection of any asymmetrical structure around a star or substellar object. While the Neyman-Pearson test is the most specific test possible, as it is set to detect a single known signature, the energy detector is the least specific test possible. As we will see, placing additional constraints on the signature $\boldsymbol{x}$ improves on the performance over T_E.

3.2.4 Signature of a binary

For this particular application, I consider the object to be a binary: a pair of "compact"[1] objects of mismatched brightness. This corresponds to a common case in astronomy, since most evolved stellar systems do not present extended features such as disks or rings of a brightness comparable to stellar companions such as large planets or other stars. Each luminous source is assumed to be unresolved, in the sense that each object making up the pair can be modelled as a point source.

For a pair of point sources of different intensities, the 2D, on-sky intensity can be modelled as a pair of Dirac functions:

$$O(x, y) = \delta(x, y) + c^{-1}\delta(x - \alpha, y - \beta). \tag{3.40}$$

The position of the companion is decomposed into its x and y components, and its coordinates are α and β. The contrast c is defined as the flux of the central object divided by the flux of the companion.

[1] Here, "compact" refers to self contained, gravitationally bound objects, such as stars or planets, rather than to white dwarfs of black holes.

The complex visibility $\mathcal{V}_0(u,v)$ associated to this object is the 2D Fourier transform of Eq. (3.40) (Cittert, 1934; Zernike, 1938) at the spatial frequency (u,v). Since the Fourier transform is linear, the complex visibility of a binary can be written as the sum of the Fourier transforms of each Dirac function.

The Fourier transform of a two dimensional function $f(x,y)$ at the spatial frequency (u,v), $F(f)(\frac{u}{\lambda},\frac{v}{\lambda})$[2] is given by

$$F(f)(\frac{v}{\lambda},\frac{v}{\lambda}) = \int\int f(x,y)e^{-2\pi i(x\frac{u}{\lambda}+y\frac{v}{\lambda})}\mathrm{d}x\mathrm{d}y, \tag{3.41}$$

where the normalisation is neglected. This normalisation factor only affects the amplitude of the Fourier transform, not its phase. For the central object, since $\delta(x,y)=0$ for $x\neq 0$ and $y\neq 0$, and $\int_{-\infty}^{\infty}\delta(x)\mathrm{d}x=1$, its Fourier transform is

$$\begin{aligned} F\Big(\delta(x,y)\Big)(u,v) &= \int\int \delta(x,y)e^{-2\pi i(x\frac{u}{\lambda}+y\frac{v}{\lambda})}\mathrm{d}x\mathrm{d}y \\ &= e^{-\frac{2\pi}{\lambda}i(0u+0v)} \\ &= 1, \end{aligned} \tag{3.42}$$

and for the companion,

$$\begin{aligned} &F\Big(c^{-1}\delta(x-\alpha,y-\beta)\Big)(u,v) = \\ &c^{-1}\int\int \delta(x-\alpha,y-\beta)e^{-\frac{2\pi}{\lambda}((x-\alpha)u+(y-\beta)v)}\mathrm{d}x\mathrm{d}y \\ &= c^{-1}e^{-\frac{2\pi}{\lambda}i(\alpha u+\beta v)}. \end{aligned} \tag{3.43}$$

The sum of these two Fourier transforms gives the complex visibility vector:

$$\boldsymbol{\mathcal{V}}(\boldsymbol{u},\boldsymbol{v}) = 1 + c^{-1}\exp\left(-i\frac{2\pi}{\lambda}(\alpha\boldsymbol{u}+\beta\boldsymbol{v})\right). \tag{3.44}$$

The resulting visibility depends on three parameters: two of them define the position of the companion relative to the brighter object, while the last one defines the contrast, or flux ratio between the two objects. Compared to the previous case, where there are as many degrees of freedom as the number of elements comprised in the kernel phase vector (typically a few hundreds), having only three parameters. Recalling that the kernel phases vector is defined as $\boldsymbol{k} := \boldsymbol{K}\boldsymbol{\phi}$ (Eq. (3.1)), and that the whitened kernel phases vector is given by the whitening of $\boldsymbol{k}$, $\boldsymbol{y} := \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{k}$ (Eq. (3.1)), since, in the absence of errors, $\boldsymbol{k} = \boldsymbol{K}\boldsymbol{\phi}_0$, we can deduce the expected value for the phase of the

[2]The "kernelific" way of managing spatial frequencies is to treat them as distances, and to normalise by the observation wavelength when modelling the object to get back to dimensionless quantities. This is different from the usual approach in interferometry, where the spatial frequencies are given as distance divided by the observation wavelength. This allows the same pupil model to be used at several wavelengths.

object's visibility

$$\boldsymbol{\phi}_0 := \angle \boldsymbol{\mathcal{V}}_0, \tag{3.45}$$

and therefore for the object's whitened kernel phase signature

$$\boldsymbol{x} := \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{K} \boldsymbol{\phi}_0. \tag{3.46}$$

This is the same signal as the one detected by the Neyman-Pearson test in Eq. (3.13), which corresponds here to knowing ρ, θ and c. The observed, whitened ker-phases are therefore given by:

$$\boldsymbol{y} = \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{K} \angle \left(1 + c^{-1} \exp\left(-i \frac{2\pi}{\lambda} (\alpha \boldsymbol{u} + \beta \boldsymbol{v})\right)\right) + \boldsymbol{\epsilon}. \tag{3.47}$$

In this case, the MLE is defined as:

$$\begin{aligned} \hat{\boldsymbol{x}} :&= \underset{\alpha,\beta,c}{\operatorname{argmax}}\, \ell(c, \alpha, \beta; \boldsymbol{y}) \\ &= \underset{\alpha,\beta,c}{\operatorname{argmax}} \exp -\frac{1}{2} \| \boldsymbol{y} - \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{K} \angle (1 + c^{-1} \mathrm{e}^{-i \frac{2\pi}{\lambda} (\alpha \boldsymbol{u} + \beta \boldsymbol{v})}) \|^2. \end{aligned} \tag{3.48}$$

Barring any linear approximation of the binary's signature (which will be examined in the final part of this chapter), there is no relation that enable the direct estimation of the MLE from the data vector $\boldsymbol{y}$. The maximisation has to be performed numerically, by exploring the parameter space to find the value of the signature that maximises the likelihood $\ell(\hat{\boldsymbol{x}}; \boldsymbol{y})$. This is a process detailed in Sec. 3.3.4.

Once the MLE has been determined, the GLR, introduced in Eq. (3.28) gives the test:

$$T_B(\boldsymbol{y}) = \frac{\exp\left(-\frac{1}{2}(\hat{\boldsymbol{x}} - \boldsymbol{y})^T (\hat{\boldsymbol{x}} - \boldsymbol{y})\right)}{\exp\left(-\frac{1}{2} \boldsymbol{y}^T \boldsymbol{y}\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{3.49}$$

Taking the logarithm of this expression and developing the content of each exponential gives the test:

$$T_{\mathrm{B}}(\boldsymbol{y}) := 2 \boldsymbol{y}^T \hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}^T \hat{\boldsymbol{x}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta\,. \tag{3.50}$$

Determining the distribution of T_{B} under $\mathcal{H}_0$ and $\mathcal{H}_1$ requires the knowledge of the distribution of the MLE $\hat{\boldsymbol{x}}$. In many cases, there is no analytical solution for the expression of $\hat{\boldsymbol{x}}$ as a function of $\boldsymbol{y}$. For this reason, the determination of the false alarm and detection rates for T_{B} necessitates the use of numerical simulations.

3.2.5 χ^2 "goodness of fit" and T_B

The test statistic T_B can be interpreted in terms of χ^2-derived intervals. Let $\hat{\boldsymbol{x}}$ be some model obtained by some fit on data $\boldsymbol{y}$. The χ^2 score corresponding to this fit is

$$T_{\chi^2}(\hat{\boldsymbol{x}}, \boldsymbol{y}) := \sum_{k=1}^{N} (\hat{\boldsymbol{x}}_k - \boldsymbol{y}_k)^2 = (\hat{\boldsymbol{x}} - \boldsymbol{y})^T (\hat{\boldsymbol{x}} - \boldsymbol{y}). \tag{3.51}$$

Considering the likelihood in Eq. (3.14), this shows that if $\boldsymbol{y}$ is Gaussian with mean $\hat{\boldsymbol{x}}$, the score in Eq. (3.51) is indeed a χ^2_p random variable. Now, the test statistics T_B can be rewritten as

$$\begin{aligned} T_B(\boldsymbol{y}) &= 2\boldsymbol{y}^T\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}^T\hat{\boldsymbol{x}} = \boldsymbol{y}^T\boldsymbol{y} - \left((\hat{\boldsymbol{x}} - \boldsymbol{y})^T(\hat{\boldsymbol{x}} - \boldsymbol{y})\right) && (3.52)\\ &= T_{\chi^2}(\boldsymbol{0}, \boldsymbol{y}) - T_{\chi^2}(\hat{\boldsymbol{x}}, \boldsymbol{y}), && (3.53) \end{aligned}$$

which shows that T_B can be interpreted as the reduction in the sum of squared residuals when comparing the null hypothesis to the considered model.

For the sake of accurately controlling the false-alarm rate, we note however that T_{χ^2} in Eq. (3.51) may not be distributed as a χ^2_p variable because $\hat{\mathbf{x}}$ is a random variable. Actually, the true distribution of T_{χ^2} may not be known analytically, and a Monte Carlo procedure (such as that mentioned in Sect. 3.2.4 for the estimation of the correspondence between the P_{FA} and threshold for T_B) is required.

3.3 Practical case: Parameters estimation, error margins, numerical methods an detection of faint brown dwarf companions

In this section, I evaluate the performance of the three detection tests presented in Eq. (3.18), (3.32) and (3.50) for the detection of companions to faint, cool brown dwarfs in the solar neighbourhood (at distances smaller than 20pc) with images from JWST. I also present a prediction of the error margins that can be expected for the measurement of the parameters of possible companions to Y type, ultracool brown dwarfs.

As long as kernel phases can be extracted from an image (which requires optical path differences smaller than a fraction of the wavelength, as seen in Sec. 2.1.3, and a field of view smaller that the isoplanetism angle), the statistical detection procedures introduced in the previous sections can be used. Here, I present a practical application on JWST images. At the time of writing, the telescope has not been launched yet, but the optical systems and detectors have been thoroughly characterised, allowing for the estimation of the performance (see for instance Oliveira et al. (2018), Greenbaum,

Sivaramakrishnan, and Pueyo (2013), Greenbaum et al. (2014), Howard and Feinberg (2009)).

The sensitivity and stability of JWST make it ideal to observe very faint sources at sensitivities inaccessible to current observatories. Among these objects, Y-type brown dwarfs are very challenging targets for high angular resolution, high contrast observations. The coronagraphs aboard JWST are not able to probe separations bellow a few λ/D (Perrin et al., 2018), and, although one imager is capable of non redundant aperture masking interferometry, this technique requires long exposure times because of the low throughput of the mask.

For these reasons, kernel phases have an important scientific potential for JWST observations, in conjunction with the statistical methods detailed above. For the study of the multiplicity of faint objects, the operational test T_B is particularly well suited, and we will explore its implementation and performance on simulated images.

3.3.1 Y type brown dwarfs and JWST observations

The scientific use case we focus on is the study of the multiplicity of Y-type brown dwarfs. Brown dwarfs are objects which are thought to have formed in a similar fashion to stars, i.e. from an interstellar cloud of gas and dust, but with far lower masses. This translates into lower pressure and temperature at their core, and either a partial ignition of nuclear fusion, or no ignition at all (Burrows et al., 2001).

Y type brown dwarfs belong to the latter type: they represent the lower tail of the mass distribution of brown dwarfs, with surface temperatures ranging from 600 to 300 K. These low surface temperatures, together with their small size make those objects very difficult to detect: the first of these objects was discovered in 2009 (Subasavage et al., 2009), and less than thirty of them are known in total (Cushing, 2019), most of them having been discovered thanks to the Wide-Field Infrared Survey Explorer, or WISE mission (Cushing et al., 2011).

Attempts were made to detect close companions around some of these dwarfs (Fontanive et al., 2018) but their very low luminosity severely limit the contrasts achievable from the ground, because of background infrared radiation from the atmosphere, and the need to resort to laser guide star AO, which provides imperfect tip-tilt correction.

These observational difficulties will however be greatly alleviated by JWST, which will provide high angular resolution imaging in the mid-infrared, where Y-type brown dwarfs are the brightest. For this use case, kernel phases are an interesting set of observables. Furthermore, the mechanical stability of places the complex visibilities extracted from its images well within the linear regime the kernel method takes advantage of.

Using simulated JWST images, we evaluate the statistical tests introduced in the previous sections, as to provide both the demonstration of a real world

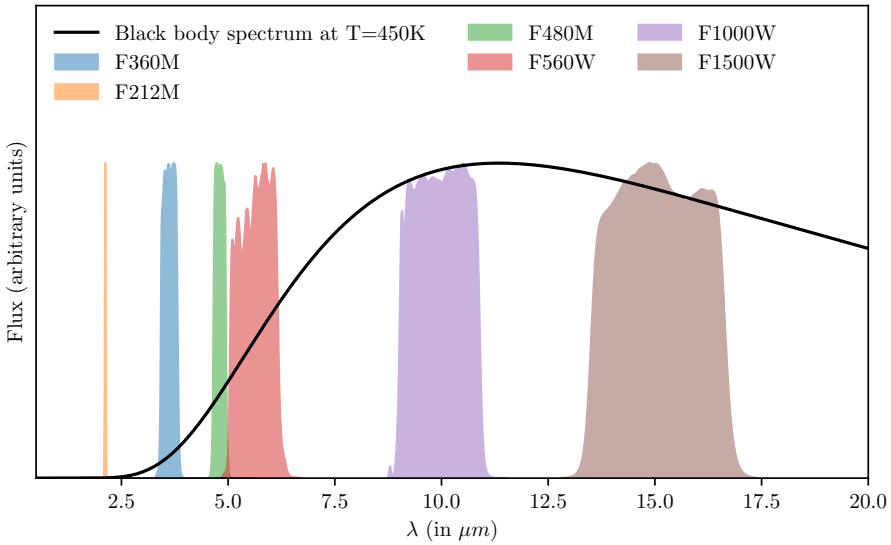

FIGURE 3.4: The spectrum of a black body at T_{eff} = 450K, overlaid with the transmission curves of some filters available for imaging with JWST. Even though observing in the close infrared provides a higher resolution, Y dwarfs are very faint in these wavelengths. Filter transmissions retrieved from *NIRCam Filters* (2019), *NIRISS Filters* (2019) and *MIRI Filters and Dispersers*.

use case for these tests, and contrast detection performance estimates for the study of Y-type brown dwarfs, opening the way for future observations.

JWST has three instruments capable of imaging: the near infrared camera (NIRCAM), the fine guidance sensor / near infrared imagers and slitless spectrograph (NIRISS), and the mid-infrared imager (MIRI). Thanks to an array of overlapping filters, these three instrument provide continuous wavelength coverage from 600 nm to 28.8 μm, with NIRCAM providing imaging between 600 nm and 5 μm, MIRI between 5 μm and 28.8 μm, and NIRISS between 600 nm and 5 μm. In Fig. 3.4, we present the transmission curves of filters that provide imaging with a sampling sufficient for kernel analysis, overlaid with the spectrum of a black body at T=450K, a typical temperature for a Y type dwarf. As we can see, these objects are very faint in the near infrared, and using wavelengths shorter that around 2.5μ m would yield an extremely low flux.

For the detection of companions near Y dwarfs, we investigate a case where images are taken using NIRISS with the F480M filter. This is motivated by the fact this wavelength domain offers a compromise between maximising the flux and the resolution. The filter's bandwidth is also narrow enough not to create important chromatic mismodellings. Using the longest wavelength available with this camera also alleviates sampling constraints: for shorter observing wavelengths, the images do not fulfils the sampling requirements,

since the pixel pitch is greater than $\lambda/2D$.[3] This constraint means that, given the platescale of the instruments, the wavelengths covered by the various filters and the diameter of the entrance aperture, the images exploitable for kernel analysis are:

- for NIRCam in the short wavelength channel (0.6 – 2.3 μm), with a platescale of 31 mas/pixel: F212N
- for NIRCam in the long wavelength channel (2.4 – 5.0 μm), with a platescale of 63 mas/pixel: F430M, F460M, F466N, F470N and F480M.
- for NIRISS, with a platescale of 65 mas/pixel (STSCI, 2018): F430M and F480M.
- for MIRI, with a platescale of 110 mas/pixel: all filters but F770W and F780W.

In the F480M band, images are simulated for 80 minutes exposures, equally split between a calibrator and the observed object of equal brightness. The package `ami_sim`, itself based on `webbpsf`[4] was used for the simulation, with a first run including no drift in the optical path difference between the calibration frame and the on target frame. The exposures are simulated for targets of luminosities corresponding to WISE 1405+5534 (W2 mag = 14.1) and WISE 0359+5401 (W2 mag = 15.4): the levels and distribution of the noises are adjusted to match those that would be experienced for actual images of these targets (Cutri et al., 2015).

The entrance aperture, together with the pupil model used for the extraction of kernels are shown in Fig. 3.5. The array of 18 hexagonal mirrors making up the segmented primary mirror which is to be folded to fit into the launcher's fairing is visible, and leads to the distinctive shape of the outer edge of the aperture. Looking closely, the small gaps between the segments are also visible. Two of the three thin struts that hold the secondary mirror in place can be seen on top. The CLEARP mask is also visible, creating a round central obstruction as well as three thick spider arms. This mask serves to obstruct the primary focus of the telescope for mid infrared observations, in order to avoid the contamination of images with stray light that would degrade the sensitivity of the instrument.

The covariance of the statistical errors affecting the calibrated kernel phases extracted from these images is obtained thanks to our knowledge of the characteristics of the detector, as listed in Table 3.1. It is estimated without considering possible variations of systematic errors, and only takes the pixel-independent noises discussed in the previous chapter into account. Calibration errors are discussed later in Sec. 3.3.2. Once the covariance has

[3] If this criterion is not fulfilled, the larger spatial frequencies suffer from aliasing, and become unexploitable.

[4] `ami_sim` is available at https://github.com/agreenbaum/ami_sim

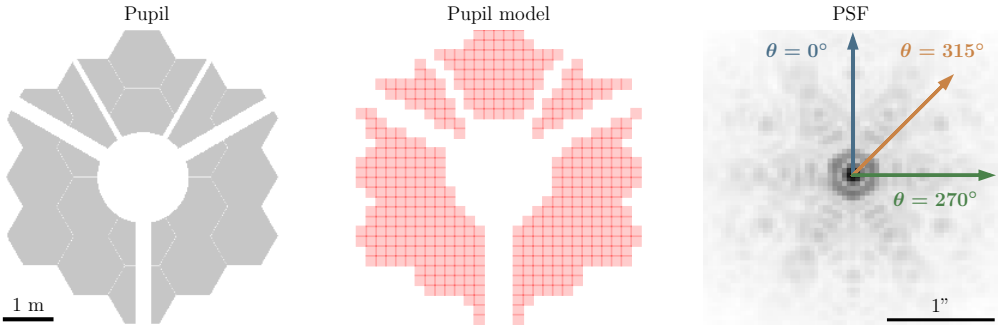

FIGURE 3.5: From left to right: The entrance pupil for JWST NIRISS using the F480M band (centred on 4.8 μm), the pupil model used to describe this pupil, and the PSF created by the pupil, together with the directions corresponding to position angles $\theta = 0°$, $\theta = 315°$ and $\theta = 270°$ considered for the simulations.

Read noise (e^-)	14.849
Flat field error	0.01%
Dark current (e^-/s)	0.04
Total integration time (s)	2400
Number of frames	15
Gain (e^-/ADU)	1.00
Jitter value (mas)	7.0
Number of photons (W2 mag = 15.4)	3.723×10^6
Number of photons (W2 mag = 14.1)	1.1181×10^7

TABLE 3.1: Detector and targets characteristics used to compute the covariance of the ker-phases extracted from our JWST NIRISS simulated dataset.

been estimated, and a whitening matrix computed, the statistical tools presented in the beginning of this chapter can be deployed.

3.3.2 Kernel phases errors estimation on JWST NIRISS

Kernel phases extracted from JWST NIRISS images are affected by errors, as discussed in the previous chapter (see Sec. 2.3). The estimation of statistical, detector errors is rather straightforward: using the exposure parameters listed in Table 3.1, it is possible to determine the distribution of the errors that affect each pixel from the image by knowing the characteristics of the detector, which have been characterised (Doyon et al., 2012). Using the obtained image as a basis, it is then possible to deduce the covariance of the noise that affects the kernel phases using Monte Carlo simulations of as many frames as necessary. Systematic, calibration errors are more difficult to predict. They are expected to stem mostly from calibration drifts, where change in the aberrations between the calibration frame and the science frame create unknown kernel phase residuals.

On JWST, wavefront drift is limited by the extreme stability of the instrument (on top of its effects being strongly reduced by the inherent robustness of kernel phases). In the past decades, stability has emerged as a primordial requirement for high resolution instruments, instead of considering having reached the diffraction limit[5] to be "good enough“. As an example of old attitudes towards calibration, the calibration and target frames on the 1997 HST observations analysed by Laugier et al. (2019b) using kernel phases were acquired weeks apart. Over this interval, significant optical drifts introduce important calibration residuals. As a result, the ker-phases extracted for Gliese 494 are dominated by calibration residuals, even after kernel treatment[6].

With a proper calibration approach, the wavefront drift between the acquisitions of the calibration frame and of the science frame can be very low. The main origin of wavefront drift is thermal: since JWST is an infrared telescope, and will work at wavelengths of up to 28.8 μ m, it needs to stay cold. This is achieved by the large sunscreen which shields the primary and secondary mirrors, as well as some of the instruments. This creates a steep temperature gradient between the cold and warm side. When the slew angle of the telescope changes, the angle of incidence of the sun on the sunshield changes, leading to a variation in the total flux received. This in turns affects the thermal equilibrium. Since the telescope is in space, thermal equilibrium

[5]This limit is in turn a somewhat arbitrary limit, usually defined as the Strehl ratio being higher than a set threshold.

[6]This does not in any way mean that kernel is worthless in this case; On the contrary, kernel treatment in Laugier et al. (2019b) have revealed a companion to Gl 494 that had not been previously detected in the 1997 archival data, and it is indeed the oldest dataset in which the companion could be resolved

is reached purely radiatively. This leads to very long thermal relaxation times, of the order of weeks (Johnston et al., 2004).

Thermal wavefront drifts have been modelled by Perrin et al. (2018). Table 1 in this paper summarises the worst-case wavefront drifts on different timescales. Over 2 hours, a timescale we can hope to reach for pairs of 40 minutes exposures, the worst-case wavefront drift is expected to be on the order of 16 nm. Translating this error value into ker-phase residuals is difficult without access to advanced simulation tools, but the residual can nonetheless be estimated: the simulation tool `webbpsf` comes with 10 "standard" OPD maps of the pupils (or phasecreens). These phasescreens are normalised, and are used to simulate the images that serve to estimate the ker-phase residuals. The standard deviation thus estimated is then used to "load" the diagonal of the covariance matrix.[7]

3.3.3 Errors on the measured parameters

With a covariance estimated, it is possible to determine the MLE expected from a dataset acquired under certain observing conditions. Here, we show the errors on the measured parameters of binaries observed under the conditions listed in Table 3.1, where the errors on measured kernel phases are dominated by image noise. I demonstrate how simply reporting error bars may not be sufficient, as their exist complex correlations between the measured contrast and separation of a companion.

The most adapted to the detection of companions of the three tests T_{NP}, T_E and T_B is T_B, since it offers a performance improvement over T_E, but does not require prior knowledge of the parameters of the researched binary like T_{NP}. To implement the test T_B, it is necessary to determine the MLE for a given dataset $\boldsymbol{y}$ (cf Eq. (3.48)). This is equivalent to estimating the parameters (α, β, c) of the binary. Since the data is noisy, the estimations of these parameters are affected by errors, which directly impact the performance of T_B: here, we present a set of methods to estimate them, and report the predicted error margins for JWST NIRISS observation of faint targets. As a first step, let us focus on the errors affecting the measured position and contrast for a detected companion's parameters. We assume the numerical estimate of the MLE $\hat{\boldsymbol{x}}$ unbiased and exact.

Using T_B first requires estimating the MLE $\hat{\boldsymbol{x}}$, i.e. the parameters of the binary which "best explain" the observed ker-phase $\boldsymbol{y}$. Any data vector leads to a MLE, whether a detection occurs or not. The uncertainty on the measured parameters is worth knowing, as the determination of the position

[7] The non diagonal terms are hard to estimate from this few realisations, and a $p \times p$ with $p \approx 1000$ covariance matrix obtained with only ten images cannot be inverted. In the following, we use either the covariance estimated in the absence of significant wavefront drift cased errors (in which case, no wavefront drift is specified), or with a covariance that is estimated by taking into account both the contributions of the noise and of the drift.

and contrast of a companion is as precious an information as the detection itself.

While in Sec. 3.2.4, the position of the companion was expressed in Cartesian coordinates to make the visibility computation easier, as uv coordinates are also expressed in a Cartesian frame of reference, here, we adopt polar coordinates, expressing the position in terms of separation from the primary and of position angle form the North axis to match the manner in which binary parameters are usually expressed. Here, we take the North axis to correspond to the direction of the coordinate β, and the East direction to the direction of the coordinate α. As a convention, let us assume the North direction corresponds to the vertical in the image. Then, the polar coordinates of a binary, consisting of a position angle θ and of a separation ρ can be found from the coordinates α and β using the relations:

$$\begin{aligned} \rho &= \sqrt{\alpha^2 + \beta^2} \\ \theta &= \arctan \frac{\beta}{\alpha}. \end{aligned} \tag{3.54}$$

I describe two methods to estimate errors, one that is adapted to the prediction of the distribution of parameters for a given noise covariance, and another one that is adapted to the estimation of confidence intervals from a dataset composed of a single realisation of $\boldsymbol{y}$

- A possible approach to *predict* the distribution of the measured parameters is a Monte Carlo simulation, involving the estimations of MLEs for a vast number of simulated realisations of $\boldsymbol{y}$. This provides a prediction of the probability that the estimated parameters under a given set of conditions fall into a given region in the parameter space. Since we are working with simulations, we estimate the error on the MLE using this method, by considering a vast number of realisations of $\boldsymbol{y}$, while keeping the signature $\boldsymbol{x}$ the same. Many realisations of the noise are then simulated to obtain a population of data vectors $\boldsymbol{y}$. For every realisation, the MLE $\hat{\boldsymbol{x}}$ is computed, giving the MLEs $\hat{\rho}$, $\hat{\theta}$ and $\hat{c}$ for each noise realisation.

  The results of this approach are show in Fig. 3.6, where it is made evident that the estimated separations and contrasts are linked: the plot of one as a function of the other has a banana shape, especially for relatively high contrasts and low separations, where the power of the signal $\boldsymbol{x}$ becomes small relative to the power of the noise $\boldsymbol{\epsilon}$. Even if the contrast and the separations are not well determined, the position angle can be determined with a good accuracy. Also note that, if the contrast is known from previous observations, accuracy on the error in the contrast can be greatly improved. For instance, if the contrast has been determined from previous observations (for instance, at a point

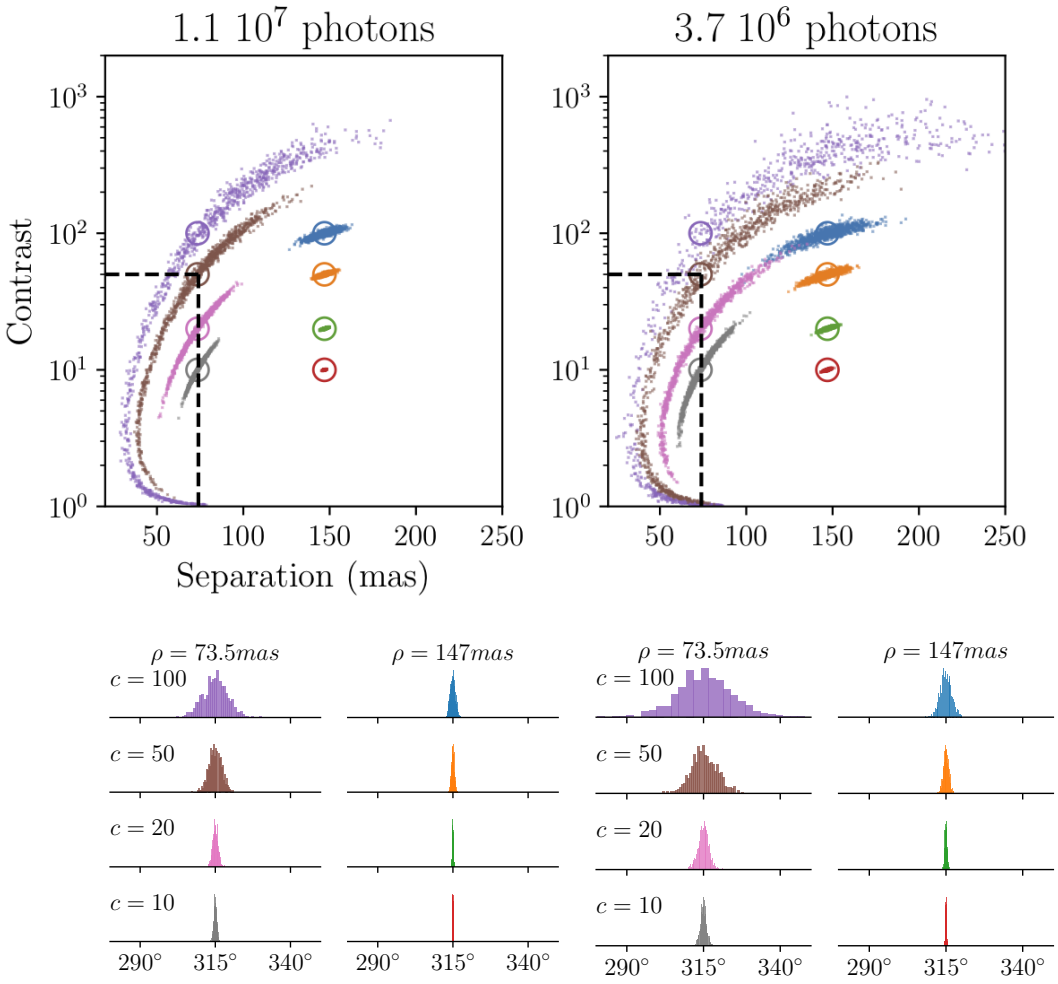

FIGURE 3.6: Error on the recovered parameters, determined by drawing many realisations of $\boldsymbol{y}$. Top panels: the circles represent the separation and contrasts of the injected signature, and each column represent one flux regime. The bottom panels represent the error on the estimated position angle θ. The parameters ρ and c of each injected signature are represented as coloured circles on the top panels, while θ is fixed at 315° for all signatures. The same colour code is used for every panel, with each colour corresponding to an individual injected signature. Each dot on the top panel represents the parameters estimated for a single realisation of the noise.

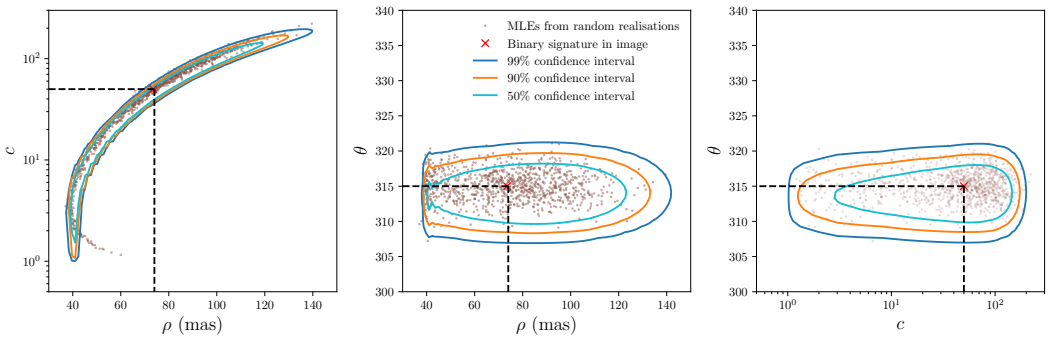

FIGURE 3.7: Isocontour map of the likelihood $\ell(\rho, \theta, c; \boldsymbol{y})$ with binary signatures $\boldsymbol{z}$ computed for a 3D grid of values taken by ρ, θ and c. The image from which $\boldsymbol{y}$ is determined has a binary signature at $\rho = 74$ mas, $\theta = 315°$ and $c = 50$. The flux is 1.1 10^7 photons.

in an orbit where the companion is farther from the main source) to be equal to say, 50, then the uncertainty on the separation in much reduced.

- If observers only have access to a single realisation, as would be the case in practice, another method can be used to determine the errors on the measured parameters, using a map of the likelihood $\ell(\rho, \theta, c; \boldsymbol{y})$ over the parameter space. By computing likelihoods over a grid of values of ρ, θ and c, if this likelihood is normalised so that its sum over all the considered parameters is 1 (this corresponds to assuming any companion would exist in the interval of parameters for which the likelihoods are calculated), confidence intervals can be drawn using the sum of the likelihood inside a given region of the parameter space. For instance, the region over which the sum of the likelihood is equal to 0.99 corresponds to the estimated 99% confidence interval.

  Fig. 3.7 shows the probability mass distribution for a binary at $\rho = 74$ mas, $\theta = 315°$ and $c = 50$, in the flux regime shown on the left had side of Fig. 3.6, at 1.1 10^7 photons reaching the detector during a 40 minutes exposure. The errors obtained using a the MC simulations are also shown, and match the confidence intervals. This sort of representation is more appropriate for reporting error bars in practice, since it can do so from a single measurement, and also provides numerical estimates for the confidence intervals.

Whatever the method to estimate the errors may be, the representation of the error bars is crucial: simply giving error *values* on the recovered parameters of a binary does not show the full picture, as it often assumes the region in which the parameters lie has the shape of an ellipsoid. In reality, these regions can take complex shapes, which are not necessarily captured by simple errors bars.

For the three parameters that determine the signature of a binary, a grid search is not computationally prohibitive for the determination of a *single* likelihood map: it consists of $N_\rho \times N_\theta \times N_c$ products of vectors of length p, with N the grid size, and p the number of ker-phases. However, the determination of the distribution of the test statistic of T_B under the null hypothesis and, more importantly under $\mathcal{H}_1$ for all the parameters one wishes to include in the detection limits map requires the computations of many MLEs, for many realisations of $\boldsymbol{y}$. This means a systematic exploration of the parameter space is not ideal in those circumstances, and that faster numerical methods are required.

3.3.4 Numerical MLE estimation.

Since estimating the MLE with a grid search is not feasible for a big number of realisations in a reasonable time, determining the detection probability and the false alarm rate for a given signature $\boldsymbol{x}$ requires an efficient and reliable way of determining $\hat{\boldsymbol{x}}$ under both $\mathcal{H}_0$ and $\mathcal{H}_1$.

The one we ended up using in practice consists in using the injected signature as a starting point for a gradient descent procedure.[8] This method presents the advantage of being very quick: only a handfuls of likelihoods need to be computed to find the MLE for a given realisation. Although it works very well at favourable signal to noise ratios,in which case the valley of global minimum is very wide and deep, it is sensitive to local minima, and to producing spurious MLEs if the signal-to-noise ration is bad enough. The application of this method becomes more complicated under $\mathcal{H}_0$, where it is difficult to find a starting point in the parameter space. In practice, one could therefore use either a starting point at zero separation and at an arbitrary contrast, or a random starting point, in the bounds of the separations accessible to the aperture model.

Alternatively, one can use an exhaustive search over a grid, where a likelihood is computed at each point of a grid in the parameter space. Depending on the definition of the grid, a gradient descent method can then be employed to hone in on the exact MLE. This method runs little risk of giving the wrong MLE due to a local minimum, but it is computationally expensive. We have found that in practice, when it comes to separations, the grid step cannot be smaller than around one half λ/D. With an aperture model that enables the measurement of kernel phases of binary at up to 10 λ/D, the grid for the positions must be at least 40×40. On top of this, the contrast needs to be determined. A $40 \times 40 \times 40$ grid search necessitates the computation of 64,000 distinct likelihoods. For a single measured kernel phases, it is not an extreme cost. But for the purpose of finding detection limits, which involves

[8] Although the estimation of $\hat{\boldsymbol{x}}$ relies on maximising the likelihood, the equivalent approach of minimising the opposite of the term in the exponential in Eq. (3.27) is more practical to implement.

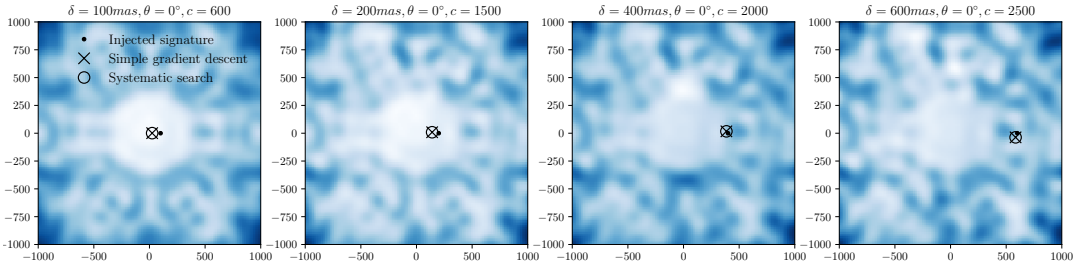

FIGURE 3.8: For contrasts above the detection limits (see Fig. 3.11 in the higher flux case), the MLE estimated by gradient descent (represented with a cross) closely matches the likelihood estimated with an exhaustive search (represented by a circle). This means that the estimation of the distribution of the test statistic under $\mathcal{H}_1$ is unbiased by the use of a simple gradient descent. As the separation increases, the width of the valley in which the minimum is situated decrease, however, using a gradient descent with the parameters of the injected signature as a starting point remains reliable.

hundreds or thousands of evaluations at each point of a grid, this high cost makes the method prohibitive.

However, for the purpose of determining the MLE from a single, real observation, a grid search is tractable, and poses no risk of falling into a local maximum. For these reasons, it is the preferred method in this case. For the purpose of determining detection limits however, speed is much more of concern, and a gradient descent with the parameters of the injected signature as a starting point would be preferred. This choice is further validated by the fact that the regime in which the multimodal nature of the likelihood becomes an obstacle to the determination of the MLE falls outside of the detection limits. This is shown in Fig. 3.8.

3.4 Performance

The contrast detection limit for binaries in images depend on the performances of the detection tests introduced before. In this section, we compare the general performances of the three tests in Sec. 3.4, predict the mass detection limits for a known Y type brown dwarf in Sec. 3.4.2, and look at the maximum attainable contrast performance in Sec. 3.4.3.

3.4.1 Detection and contrast performance

Firstly, we validate the theoretical relations predicting the performance of the NP test T_{NP} (Eq. (3.25)) and of the energy detector T_E (Eq. (3.39)), and determine the actual performance of T_B (Eq. (3.50)). For that purpose, we

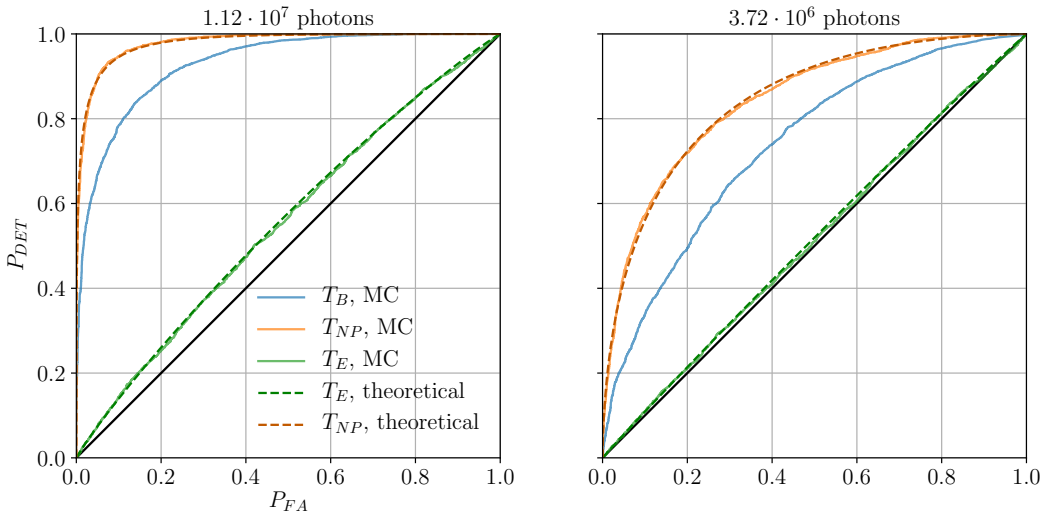

FIGURE 3.9: ROC curves of T_E (green), T_{NP} (blue), and T_B (orange). Theoretical ROC curves for T_{NP} and T_E plotted using Eq. (3.25) and Eq. (3.39), for a companion at $\rho = 200$ mas, $c = 1200$, and $\theta = 45°$ off the vertical. Dashed lines correspond to theoretical ROCs, while solid lines represent ROCs obtained by Monte-Carlo simulations. The closer a curve is to the black line on the diagonal, the less powerful the corresponding test. The higher flux regime is represented in the top panel, and the lower flux regime in the bottom panel. The performance of T_{NP} and T_E are accurately described by the theoretical expressions in Eq. (3.25) and Eq. (3.39). The test T_{NP} presents the highest performance. T_B is the next-best-performing test and T_E has the lowest performance of the three. We see a clear improvement of the power of all tests as the flux (and thus the S/N) increases.

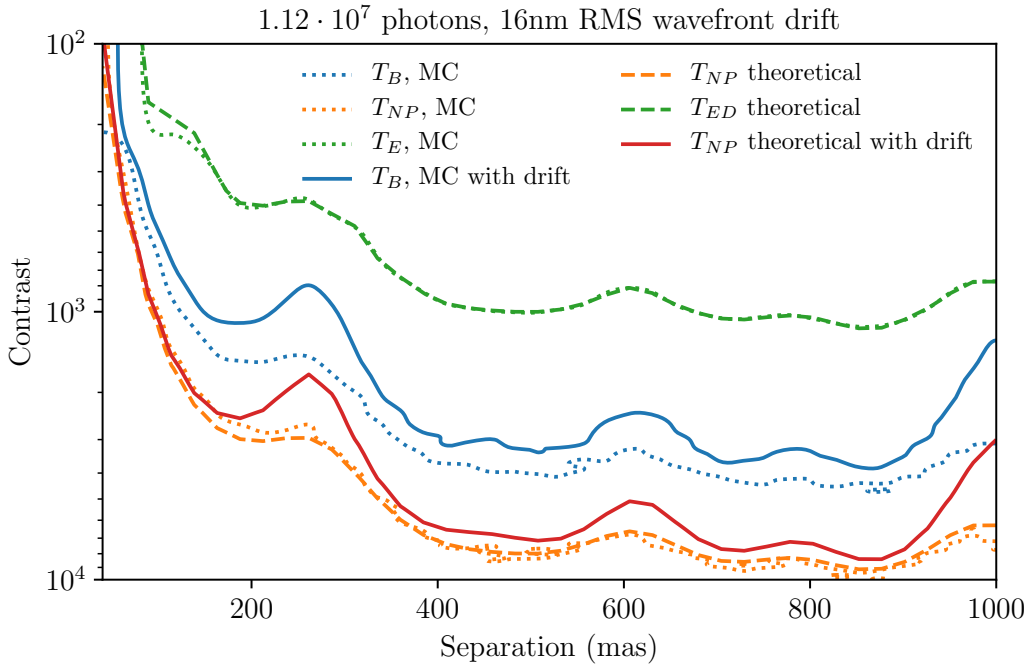

FIGURE 3.10: Detection limits at a fixed position angle $\theta = 315°$ (the isocontours below which the P_{DET} falls below 68% for a fixed P_{FA} of 1%), represented as a function of the separation and contrast of the companions, for T_E (green), T_B (blue), and T_{NP} (orange) The dashed lines represent theoretical detection limits for T_E and T_{NP} (Eq. (3.25) and Eq. (3.39)) and the dotted lines present the limits achieved in the MC simulations. T_{NP} (orange) provides ideal detection limits for a Kernel treatment of a JWST-NIRISS image and the practical test T_B (dotted blue) has contrast detection limits within a factor 2.5 of the theoretical maximum. The solid lines represent the detection limits for T_B (blue) and T_{NP} (red) with a calibration residual corresponding to a 16 nm RMS wavefront drift.

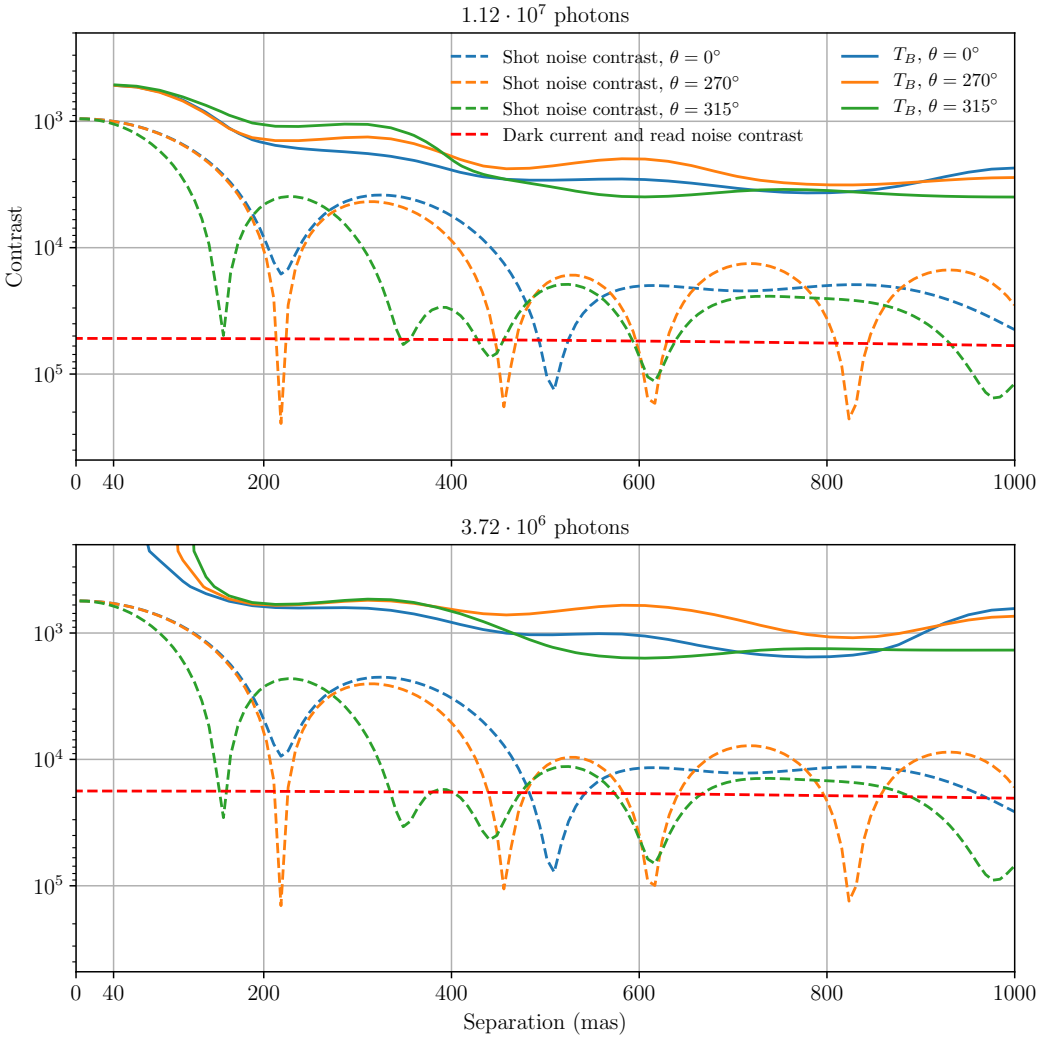

FIGURE 3.11: Detection limits for test T_B (Eq. (3.50)), in the higher flux regime (top panel) and the lower flux regime (bottom panel). The solid lines correspond to contours of P_{DET} = 68% at a fixed P_{FA} = 1%. Detection limits are represented at three different position angles for the companion: 0, 45, and 90 degrees off the vertical, as orientated in the PSF shown in Fig. 3.5. The relative S/N (see text) are indicated by dashed lines. The shot (photon) noise is the main limiting noise in most cases.

perform Monte Carlo simulations consisting of 2000 realisations[9] of $\boldsymbol{y}$ under $\mathcal{H}_0$ and under $\mathcal{H}_1$ for a given signature $\boldsymbol{x}$ (cf Eq. (3.44)).

All of the detection limits are shown for $P_{FA} = 1\%$ and $P_{DET} = 68\%$. In terms more frequently encountered in astronomy publications, this is equivalent to having a 68% chance of making a $\approx 2.3\sigma$ detection, as the distribution function of the one dimensional, normal distribution $\mathcal{F}_{\mathcal{N}(0,\sigma)}(\xi)$ is equal to 0.01 for $\xi \approx 2.3\sigma$.

On each realisation of $\boldsymbol{y}$, we perform each of the three tests using the kernels operator $\boldsymbol{K}$ and the covariance matrix $\boldsymbol{\Sigma}$. Figure 3.9 presents our results in the form of ROC curves, which provide a graphical representation of the power of each test. It can be seen that the dashed curves representing the theoretical ROCs accurately match the solid lines corresponding to the performance achieved in practice. As expected, T_{NP} appears to be the most powerful of the three tests (this test corresponds to the upper performance bound) and T_E the least powerful of the three (this test uses no prior information on the target signature and can be seen as a lower bound). The performance of T_B logically lies in between, but much closer to the upper than to the lower bound.

The detection limits for the three tests T_{NP}, T_E, and T_B are represented in Fig. 3.10 across a range of contrasts and separations, for a fixed position angle $\theta = 315°$. The dashed lines correspond to no wavefront error while the solid lines correspond to 16 ñm RMS of wavefront error. We can see that the theoretical performance, validated for a single companion signature in Fig. 3.9, hold true over a large range of contrasts and separations, and that the detection limit of T_B remains close to the bound provided by T_{NP}. The dashed and dotted lines correspond to a perfectly stable JWST leading to a perfect calibration of the systematic errors.

The detection limits further depend on θ, because the PSF of JWST NIRISS is not centrosymmetric (as visible in Fig. 3.5). Fluctuations of these limits are shown in Fig. 3.11 for three position angles. Figure 9 also indicates the S/N level at the corresponding positions in the image (computed here as the maximal pixel value of a noiseless image with only the companion, divided by the standard deviation of the considered noise), showing that the detection limits follow the overall noise level in the image. Performance wise, the detectable contrast ratios are of the order of 10^3 at 200 mas, with some variations between the two flux levels considered.

3.4.2 Mass limits for WISE 1405+5534

WISE 1405+5534 is a Y-type brown dwarf with a W2 magnitude of 14.1 that was used as a reference target to produce the contrast detection limits featured in Fig. 3.10. The raw observational detection limit curve of contrast

[9] The number of realisations is dictated by the target P_{FA} and P_{DET}. For the considered $P_{FA} = 1\%$ and $P_{DET} = 68\%$, 2000 realisations correctly sample the distributions of the test statistic of T_B under $\mathcal{H}_0$ and $\mathcal{H}_1$.

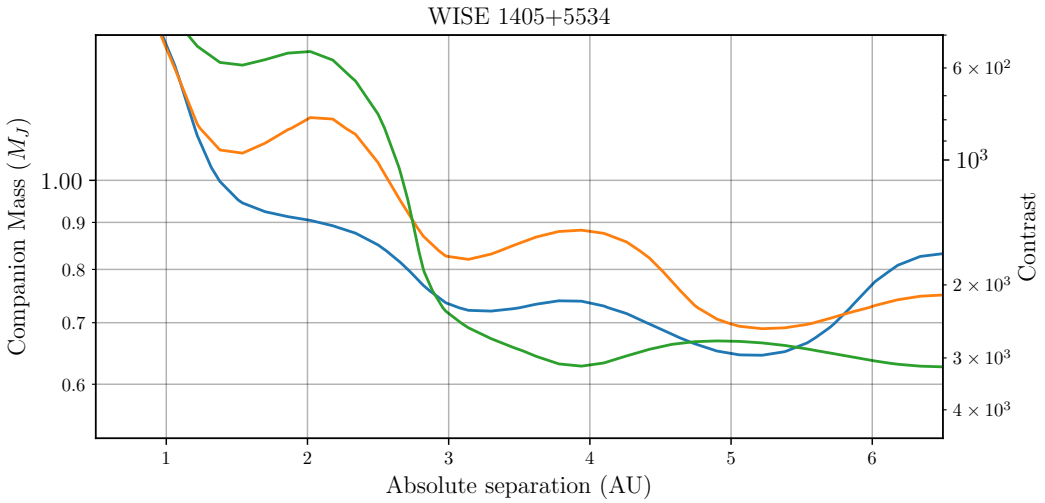

FIGURE 3.12: Detection limits of a possible companion to WISE 1405+5534 at $P_{FA} = 1\%$ and $P_{DET} = 68\%$, as a function of contrast (right ordinate axis) or mass (left ordinate axis) and absolute separation in AU. A one-Jupiter-mass object is detectable down to 1.5 AU from the primary.

as a function of angular separation can be converted into an astrophysical detection limit curve of companion mass as a function of orbital separation.

Whereas the 129 ± 19 mas parallax measured by Dupuy and Kraus (2013) directly allows for the conversion of the angular separation into a projected orbital distance, the contrast to mass conversion requires a model. We use the mass–luminosity relations given by the AMES-Cond model of Baraffe et al. (2003) for an age of 1 Gyr and a mass estimate of 30 M_J for the primary given by Cushing et al. (2011).

The detection limits obtained for WISE 1405+5534 are shown in Fig. 3.12. At $P_{FA} = 1\%$, and $P_{DET} = 68\%$, a $1M_J$ can be detected at separations greater that 1.5 AU. An orbit with this semi major axis would have a period of 40 years, thus a quarter of an orbit could be captured with repeated observations over the expected service life of JWST.

3.4.3 Bright limits

For the faint Y-dwarf targets considered thus far, it may have occurred to the reader that the contrast detection limits are dominated by the effect of the dark current and the readout noise and not by the photon noise of the central object. We wish here to complete the description of the properties of our approach with a bright target scenario that will feature a different behaviour, thus exhibiting the contribution of the photon noise.

The saturation limit for full-pupil JWST NIRISS using the F480M filter and a 64×64 pixels subarray size is 7.6 mag. We consider a shorter observation sequence, with a total of 20 minutes spent on the target of interest and

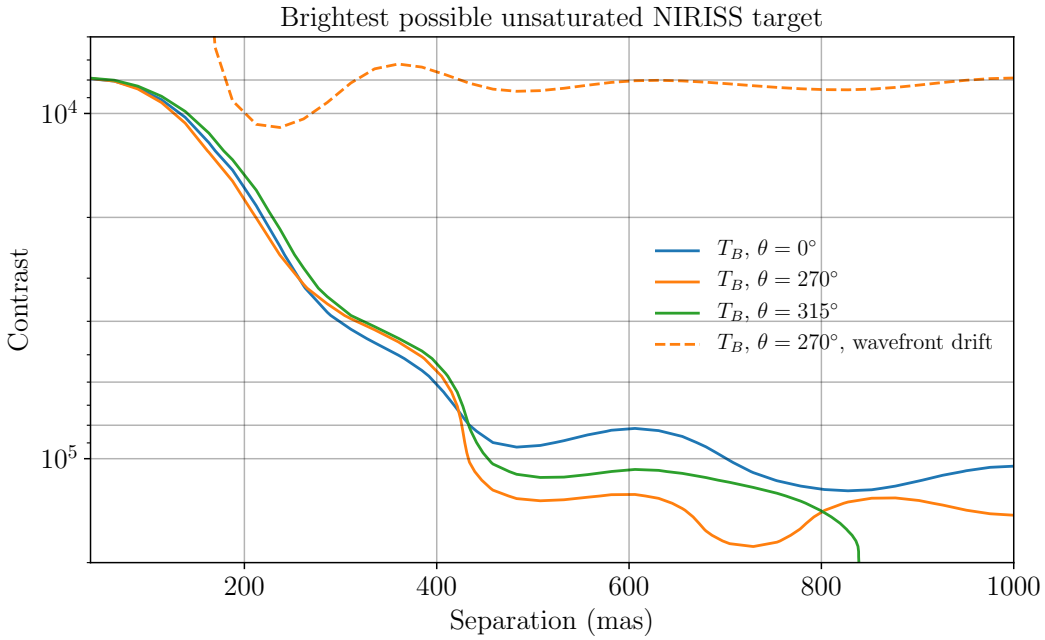

FIGURE 3.13: Detection limits for the brightest target observable without saturation with JWST NIRISS. Solid lines show detection limits for T_B at P_{FA} = 1% and P_{DET} = 68% applied to the image with the greatest possible dynamic range, with 20 minutes total integration time. For the brightest images, the kernel method with the test T_B ideally allows detection of contrasts up to 10^5 beyond 500 mas. The dashed orange line represents the more realistic detection limits in the presence of a 16 nm wavefront drift.

20 minutes on a calibrator of similar brightness. The detection limits for this observation using the operational test T_B are shown in Fig. 3.13, at $P_{FA} = 1\%$, and $P_{DET} = 68\%$.

Unlike the contrast detection limits obtained on the faint targets, the curves now clearly reveal two different regimes. Up to an angular separation of ≈500 mas, where the photon noise is expected to dominate, the contrast detection decreases as a function of the angular separation. Beyond this point, it reaches a plateau, as the detection is once again dominated by the homogeneous properties of the dark current and the readout noise.

In this bright scenario, calibration errors induced by a drift comparable to what was described in Sect. 3.4 will have a stronger impact on the weak signal of a high-contrast companion. Sallum and Skemer (2019) feature contrast detection limits for NIRCam in a similar scenario that takes calibration errors into account. Under the hypothesis introduced in Sect. 3.3.3, the calibration error accounts here for 85% of the total noise variance of the kernels and therefore results in a degraded performance by a factor of approximately 10, as shown by the dashed curve in Fig. 3.13.

In Fig. 3.6 and Fig. 3.7, it is apparent that ker-phases become less sensitive at resolutions smaller than λ/D: even if the contrast remains the same, the errors on the parameters of the binary increase at small separations. This is corroborated by a quick degradation of contrast detection limits at close separations, seen in Fig. 3.10 and all other detection maps. This does not seem to be a pure signal-to-noise ratio issue in the raw phase: even signatures with small contrasts are affected. Binary signatures seem to present a fundamental resolution limit. Poor detection performances at very small separations are not particular to JWST kernel phases either: they have been observed on other telescopes, with both ker-phases and NRM closure phases. The limits can be explained as the apparition of a fundamental degeneracy between the tip-tilt aberration mode and a companion signature at very close separations. The degeneracy appears in the first elements of the Taylor series expansion of the phase of the complex visibility of a binary, which correspond to a component identical to a “slope“ in the Fourier domain, as shown in Appendix C.

3.5 Conclusion

In this chapter, we have shown how the presence or absence of an asymmetrical structure around a star or, in this case a brown dwarf can be expressed as a set of hypotheses on the distribution of the kernel phase extracted from its image. These hypotheses can be used to design a detection test. If the signature one attempts to detect is known in advance, there exists an optimal detection test, the likelihood ratio. If it is not the case, the likelihood ratio can be generalised to the case where the signature is estimated using the measured signal, to create a generalised likelihood ratio test, or GLR.

To better illustrate the performance of of the operational test T_B, let us go back to the example in Fig. 3.1, the Fig. 3.14 shows how a companion can

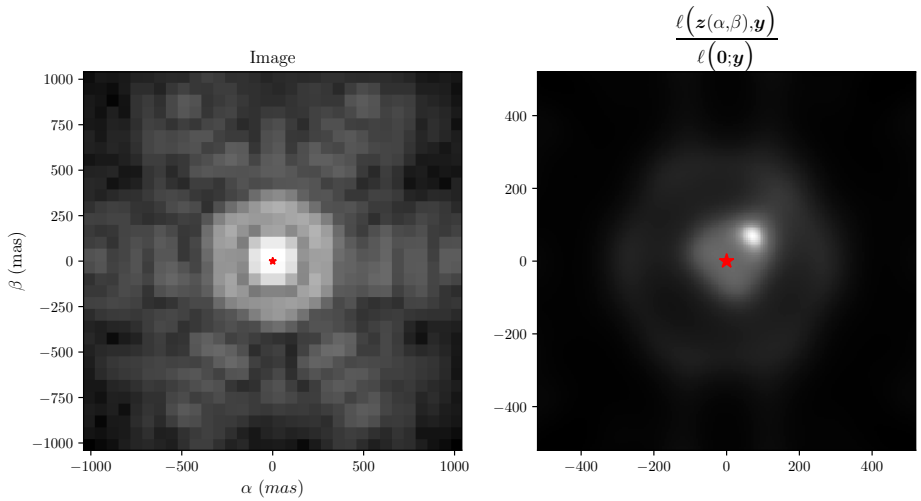

FIGURE 3.14: Going back to the example in Fig. 3.1, and more specifically to the example on the right, the likelihood ratios for different companion positions, and a contrast of 50 (on a non linear scale). The maximum of the ratio corresponds to the test statistic for $T_B(\boldsymbol{y})$. In this case, the value of this statistic enables us to guarantee $P_{FA} < 10^{-7}$, given the estimations of the distribution of the test statistic of T_B under $\mathcal{H}_0$

be detected on the image on the left using kernel analysis and the statistical methods introduced in this chapter, we can now detect the companion, which is located at $\rho = 73.5$ mas, at $\theta = 315°$ with a contrast $c = 50$, as shown in Fig. 3.14, with $P_{FA} > 10^{-7}$. Therefore, for this use case, kernel phase analysis combined with detection tests, and the great stability of JWST allows for the detection of very close binaries that would not be detectable otherwise, providing an additional significance level for the kernel detection. This achieved by exploiting both our physical understanding of how errors propagate to an image and minimising them thanks to the kernel method a subspace projection method that cancels the nuisance signal, and to the way errors in the image affect these observables by treating them statistically.

We have determined the test statistics of both the likelihood ratio and the GLR for two possible structure of the object subject to detection. Then, we saw how these tests perform for the detection of structures around a cold, faint brown dwarf observed with JWST. Focusing on the test tuned to detect companions of unknown position and contrasts, we saw how it could reach contrasts of around 10^3 well inside of the inner working angles of the coronagraphs expected to fly on JWST, with $P_{FA} < 10^{-2}$.

Kernel phase observations are therefore a compelling observing method for the study of the multiplicity of faint, cold brown dwarfs, especially coupled with statistical detection tests able to guarantee a false alarm rate. It is also worth noting that the statistical procedures detailed in this chapter can also be deployed for any pupil geometry.

The main limitation of these statistical techniques remain the characterisation of the distribution of calibration errors, which still dominate ground based observations. Without a good estimation of their distributions, it is difficult to establish the nature of the observed object, and badly estimated errors can lead to false detections, or detection of the wrong signature, like a disk asymmetry being confused with a companion. This is not on an issue for low flux targets observed with a stable platform, but it can become a limit, even with the very stable JWST, where calibration errors come to dominate for brighter targets. A possible answer to this issue could be better aberration propagation models, that can lower the residual for a given aberration, as shown by Martinache et al. (2020).

The work presented in this chapter have been the object of Ceau et al. (2019), published in Astronomy & Astrophysics. A proposal is currently being written for Y dwarf direct imaging using kernel phases, leveraging the power of these methods.

Conclusion

Throughout this thesis, we have explored the Kernel method, and the construction of aberration robust observables from the Fourier transform of images, using the methods introduced by Martinache (2010). This method relies on the modelling of image formation as the acquisition of interferometric observables from an array of regularly spaced pupils matching the shape of the entrance aperture. Contrary to a classical interferometer, this arrangement is highly redundant: for most spatial frequencies, the observed complex visibility is typically the result not of the combination of the light sampled by a single pair of subpupils, but of many pairs. This redundancy can be eliminated with a non redundant mask, which enables the formation of images from which closure phases, observables not affected by aberrations can be formed. These observables rely in a linear relation that mediates the propagation of aberration from the subpupils to the observed visibilities.

Non redundant aperture masking is a successful technique, that can be deployed on large, uncorrected ground based telescopes. In recent years, extreme adaptive optics has emerged, and has proven able to eliminate most of the atmospheric aberrations for 8 metre class telescopes: images are now dominated by diffraction rather than aberrations, with optical path differences across the entrance aperture much smaller than typical close infrared wavelengths. In this regime, a linear relation can approximate the combination of the contributions of many subpupils to form the complex visibility for a single spatial frequency. A relation similar to the one that is used to produce kernel phases can be used to construct aberration-robust observables from images obtained with a pupil of any shape: kernel phases. The key to this description is a matrix relation that links the phase shift imposed at each point of the aperture corresponding to the phase of the complex visibilities for each of the spatial frequencies sampled by the array of virtual subpupils. Linear algebra then gives us the tools to eliminate this contribution by projecting the phase of the complex visibilities in a subspace unaffected by this linear component.

In the low aberration regimes unlocked by AO and space-based observatories, kernel phases can eliminate most of the aberration signal from the phase of the complex visibility. The method is however not perfect: even when the linear approximation on which the method relies is verified, a residual aberration signal is generated. We examined different possible candidates to explain this residual: the improper modelling of the entrance aperture of the telescope, by a discrete array of virtual subpupils, aberration internal to the virtual subpupils and the non monochromatic nature of light. Of these three

factor, only the non monochromatic nature of light was found not to have any impact, and, for a "simple“ circular aperture with a central obstruction, we found that a better model of the pupil with an incorporation of subpupil transmissions and the filtering of the spatial frequencies that can be attributed to aberration inside of the subpupils can reduce the residual aberrations dramatically, as reported in Martinache et al. (2020), where we show how the aperture modelling can be improved to reduce the aberration residuals.

Some errors that are not aberrations also affect kernel phases: on the focal plane, detector and photon noises propagate to the kernel phases and limit their sensitivity to faint signatures. The distribution of this error needs to be understood for the kernel phases to be exploited. Even though the photon noise follows a Poisson distribution, the fact that kernel phases are the result of the combination of the pixel values of all pixels guarantees that their distribution is a multivariate normal. The covariance of this distribution can be estimated either by a Monte Carlo simulation, taking the realisations of many images at a given noise level; or by a linear approximation, using the individual variance or each pixel. The former option is what we used, since the computational overhead to simulate a few hundred thousand frames and to compute the Fourier transform for each of these frames is not a big hindrance. Once the covariance has been estimated, it serves either to compute the likelihood of an object's parameters given a data vector, or to construct a set of observables of covariance identity, meaning they are uncorrelated and of variance 1. This last option enables the computation of likelihoods to be simpler, and the kernel matrix can be altered to produce these whitened observables, so this is the one we adopted.

This determination of the distribution of the kernel phases enables the computation of the likelihood of obtaining a given kernel phase vector for a given object signature. The likelihood then serves as a basis for detection: by comparing the likelihood of the signature corresponding to a single star to the likelihood of it corresponding to some other structure, it becomes possible to detect the presence of a structure around a star. The optimal of close-to-optimal detection procedure changes whether the structure of interest is known in advance, unknown, or corresponds to a specific feature. We focused on the detection of a stellar or substellar companion, with an application to JWST NIRISS images of faint, cool brown dwarfs known as spectral class Y brown dwarfs, and demonstrate that it is possible to devise a detection procedure capable of detection at contrasts of 10^3 at separations of 200 mas with false detection probability smaller than 1%, a performance that can be matched neither by current ground based telescopes for these objects, nor by the coronagraphs available for JWST observations. These detection procedures and their application were the object of Ceau et al. (2019).

These procedures have also been included in the python kernel phase analysis package `xara`, thanks to the work of Romain Laugier, which enables users to perform detections on images using the same suite of tools that allows from the extraction of kernel phases.

Although the detection procedures presented in Chapter 3 are either optimal or close to optimal, there are many cases where their application is not very clear cut, as the distribution of the errors in the kernel phases are not necessarily known, and the method would greatly benefit of a better modelling of the distribution of aberration residuals: the current best practice is to be overly conservative with the detection thresholds to account for these poorly constrained errors, with a negative impact on performance.

The detection performances of these detection methods are currently being applied to outline a proposal for a survey of Y type dwarfs, led by Loïc Albert at iREx in Montreal.

In parallel, the kernel method has been expanded to include more use cases. Laugier et al. (2019b) has shown how saturation in images can be corrected to make the kernel treatment possible on images with regions outside of the linear range of CCDs. Martinache (2010) demonstrated a proof of concept for aberration robust observables from interferometric nulls.

The statistical methods presented here are not bound to binaries: any object signature that can be described as a function of a set of parameters can be detected. Such object may include, disks, star with jets, or systems composed of more than two unresolved sources. Provided calibration frames are available, and the errors on kernel phases are known, archive images can be exploited using these methods.

The kernel method, as applied to images (including those obtained with NRM interferometry) can also be extended, by measuring the subspace in which errors lie. This can be done by acquiring a great number of frames (which is possible with a fast camera and bright calibration source), and then using principal component analysis to build a response matrix that serves as the transfer matrix $\boldsymbol{A}$, from which a kernels matrix can be build, by taking the left null space of this thus determined structure matrix.

Appendix A

Demonstration of the Neyman-Pearson lemma

The Neyman Pearson Lemma (Neyman and Pearson, 1933) states that the most powerful test that discriminates between two hypotheses is he likelihood ratio. Here, we consider the case where the hypotheses are of presence or absence of a signature $\boldsymbol{x}$, with a noise $\boldsymbol{\epsilon}$ of known distribution. The demonstration that follows is largely inspired by Lehmann and Romano (1959).[1]

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y} = \boldsymbol{\epsilon} \\ \mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon}, \end{cases} \tag{A.1}$$

is a likelihood ratio, giving the Neyman-Pearson test:

$$T_{NP}(\boldsymbol{y}) = \frac{\ell(\boldsymbol{y}|\mathcal{H}_1)}{\ell(\boldsymbol{y}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{A.2}$$

The power of a test is defined by its detection rate P_{DET} for a given P_{FA}: the higher P_{DET}, the most powerful the test.

We define the rejection region of the null hypothesis $\mathcal{H}_0$ of a detection test $T_T(\boldsymbol{y})$ as the collection of data vectors $\boldsymbol{y}$ that trigger a detection, i.e. a rejection of the null hypothesis:

$$R_T = \left\{ \boldsymbol{y} : T_T(\boldsymbol{y}) \geq \xi \right\}. \tag{A.3}$$

The collection of data vectors that are not in the region R_T are in its complementary R_T^c, so that $R_T \cap R_T^c = \emptyset$ and $R_T \cap R_T^c = \left\{ \boldsymbol{y} \right\}$, the set containing all the possible values for $\boldsymbol{y}$. The false alarm rate for the test T_T can thus be rewritten as the probability of the test statistic falling into the rejection region under the null hypothesis:

$$P_{FA}(T_T) = P(R_T; \boldsymbol{H}_0). \tag{A.4}$$

[1] The noise has not to be normal in this case. The only condition on the distribution if the noise is that the likelihoods $\ell(\boldsymbol{y}|\mathcal{H}_1)$ and $\ell(\boldsymbol{y}|\mathcal{H}_0)$ can be determined.

For a given signature $\boldsymbol{x}$, the likelihood that this signal gives rise to a data vector $\boldsymbol{y}$, falling into the rejection region R, is

$$P(R, \boldsymbol{x}) : \int_{R_T} \ell(\boldsymbol{x}; \boldsymbol{y}). \tag{A.5}$$

This is akin to counting the number of times the data $\boldsymbol{y}$ falls into the rejection region for a given $\boldsymbol{x}$. The false alarm rate for a test T_T with the rejection region R_T is therefore

$$P_{FA}(T_T) = P(R_T; \mathbf{0}). \tag{A.6}$$

Proving that the Neyman Pearson test is uniformly most powerful is equivalent to proving it offers the lowest possible false alarm rate, while offering the largest rejection region of the null hypothesis in the presence of a signature $\boldsymbol{x}$. From Eq. A.6, this is equivalent to proving that every test T_A has, at an equivalent P_{FA} than T_{NP}

$$P(R_{NP}; \mathbf{0}) = P(R_A; \mathbf{0}), \tag{A.7}$$

a lower or equal P_{DET} than T_{NP},

$$P(R_A; \boldsymbol{x}) \leq P(R_{NP}; \boldsymbol{x}). \tag{A.8}$$

The probabilities in Eq. A.8 can be broken down into

$$\begin{aligned} P(R_{NP}; \boldsymbol{x}) &= P(R_{NP} \cap R_A) + P(R_{NP} \cap R_A^c; \boldsymbol{x}) \\ P(R_A; \boldsymbol{x}) &= P(R_A \cap R_{NP}) + P(R_A \cap R_{NP}^c; \boldsymbol{x}). \end{aligned} \tag{A.9}$$

Eq. A.8 is therefore equivalent to

$$P(R_{NP} \cap R_A^c; \boldsymbol{x}) \geq P(R_A \cap R_{NP}^c; \boldsymbol{x}), \tag{A.10}$$

and Eq. A.7 to

$$P(R_{NP} \cap R_A^c; \mathbf{0}) = P(R_A \cap R_{NP}^c; \mathbf{0}). \tag{A.11}$$

If Eq. A.8 is true whenever Eq. A.7 is verified, then the Neyman-Pearson detection test is the single most powerful test possible.

From Eq. A.5, we get

$$P(R_{NP} \cap R_A^c; \boldsymbol{x}) = \int_{R_{NP} \cap R_A^c} \ell(\boldsymbol{y}; \boldsymbol{x}) d\boldsymbol{y}, \tag{A.12}$$

and likewise for the other term of Eq. A.8

$$P(R_{NP}^c \cap R_A; \boldsymbol{x}) = \int_{R_{NP}^c \cap R_A} \ell(\boldsymbol{y}; \boldsymbol{x}) d\boldsymbol{y}. \tag{A.13}$$

Recalling Eq. A.2 and Eq A.3, the rejection region for the Neyman Pearson test is given by

$$R_{NP} = \left\{ \boldsymbol{y} : \frac{\ell(\boldsymbol{x};\boldsymbol{y})}{\ell(\boldsymbol{0};\boldsymbol{y})} \geq \xi \right\}, \tag{A.14}$$

and its complement by

$$R^c_{NP} = \left\{ \boldsymbol{y} : \frac{\ell(\boldsymbol{y};\boldsymbol{x})}{\ell(\boldsymbol{y};\boldsymbol{x})} < \xi \right\}, \tag{A.15}$$

giving us

$$\int_{R_{NP} \cap R^c_A} \ell(\boldsymbol{y};\boldsymbol{x}) d\boldsymbol{y} \geq \xi \int_{R_{NP} \cap R^c_A} \ell(\boldsymbol{y};\boldsymbol{0}) d\boldsymbol{y} \tag{A.16}$$

and

$$\int_{R_{NP} \cap R^c_A} \ell(\boldsymbol{y};\boldsymbol{x}) d\boldsymbol{y} < \xi \int_{R_{NP} \cap R^c_A} \ell(\boldsymbol{y};\boldsymbol{0}) d\boldsymbol{y}. \tag{A.17}$$

The relation Eq. A.5, together with the equal significance levels assumed for T_A and T_{NP} (Eq. A.11) give

$$\int_{R_{NP} \cap R^c_A} \ell(\boldsymbol{y};\boldsymbol{0}) d\boldsymbol{y} = \int_{R^c_{NP} \cap RA^c} \ell(\boldsymbol{y};\boldsymbol{0}) d\boldsymbol{y}. \tag{A.18}$$

Therefore, from Eq. A.18, Eq. A.16 and Eq. A.17, we have

$$\int_{R_{NP} \cap R^c_A} \ell(\boldsymbol{y};\boldsymbol{x}) d\boldsymbol{y} > \int_{R^c_{NP} \cap R_A} \ell(\boldsymbol{y};\boldsymbol{x}) d\boldsymbol{y}, \tag{A.19}$$

which is equivalent to

$$P(R_{NP} \cap R^c_A; \boldsymbol{x}) > P(R_A \cap R^c_{NP}; \boldsymbol{x}). \tag{A.20}$$

Therefore, the only test T_A that verifies Eq. A.11 and Eq. A.10 is T_{NP} itself. Any other test, i.e with any other rejection region than the one of T_{NP} must verify Eq. A.20 at any given P_{FA}, and is therefore less powerful than the Neyman-Pearson, or likelihood ratio test.

Appendix B

A hybrid method to find the MLE

To determine the MLE from a measured kernel phase, one can solve the minimisation problem for the likelihood by performing an exhaustive search of the parameter space (which is costly), or use a gradient descent algorithm to find the local maximum for the likelihood, but with a risk it does not correspond to the global one. Here, I present an approach that helps simplifying the exhaustive search method to make it less costly, and enable a fast, reliable estimation of the MLE.

Fig. 3.8 shows the position of the position of the companion, determined both by a gradient descent algorithm and by a search on a grid of positions. A systematic exploration of the parameter space can be used, but it takes time, as it requires many likelihood computations. An alternative is an hybrid method, where a systematic exploration is performed to give a rough estimate of the parameters of the binary, placing it in the appropriate region to get the gradient descent algorithm to fall into the valley of the correct global minimum. We call this method an altered grid search, or AGS, since it still relies on the location of some of the parameters of the binary on a grid, but with a reduced number of dimensions (2 instead of 3), which is beneficial to the speed with which the MLE, ans thus the test statistic of T_B under $\mathcal{H}_0$ can be determined. This is a crucial part in the use of this test, as knowing this distribution is a requirement for the determination of P_{FA} for any given detection. Furthermore, this distribution is affected by the covariance of the the errors on the kernel phases, it is therefore not possible to determine this distribution once and for all, it needs to be recomputed for every possible covariance.

The AGS is based on an approximation of the phase of the complex visibility of a binary. Recall the expression Eq. (3.44):

$$\boldsymbol{\mathcal{V}}_0(\boldsymbol{u},\boldsymbol{v}) = 1 + c^{-1}\exp\left(-i\frac{2\pi}{\lambda}(\alpha\boldsymbol{u} + \beta\boldsymbol{v})\right). \tag{B.1}$$

The visibility can be decomposed into its real and imaginary parts:

$$\boldsymbol{\mathcal{V}}_0(\boldsymbol{u},\boldsymbol{v}) = 1 + c^{-1}.\cos\Big(-\frac{2\pi}{\lambda}(\alpha\boldsymbol{u} + \beta\boldsymbol{v}\Big) + ic^{-1}.\sin\Big(-\frac{2\pi}{\lambda}(\alpha\boldsymbol{u} + \beta\boldsymbol{v})\Big), \tag{B.2}$$

The argument of a complex number $a + ib$ is given by

$$\angle(a + ib) = \arctan \frac{b}{a}, \tag{B.3}$$

which can be approximated to the first order, if $a >> b$ with

$$\angle(a + ib) \approx \frac{b}{a}. \tag{B.4}$$

This gives the phase of the object visibility:

$$\angle \boldsymbol{\mathcal{V}}_0(\boldsymbol{u}, \boldsymbol{v}) = \arctan \frac{c^{-1}.\sin\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u} + \beta \boldsymbol{v})\Big)}{1 + c^{-1}.\cos\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u} + \beta \boldsymbol{v})\Big)}. \tag{B.5}$$

For large values of c, $c^{-1}.\sin\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u}+\beta \boldsymbol{v})\Big) << 1$ and $1+c^{-1}.\cos\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u}+\beta \boldsymbol{v})\Big) \approx 1$. The phase of the visibility can therefore be approximated as

$$\angle \boldsymbol{\mathcal{V}}_0(\boldsymbol{u}, \boldsymbol{v}) \approx c^{-1}.\sin\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u} + \beta \boldsymbol{v})\Big), \tag{B.6}$$

therefore,

$$\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{K} \angle \boldsymbol{\mathcal{V}}_0 \approx c^{-1} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{K} \sin\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u} + \beta \boldsymbol{v})\Big). \tag{B.7}$$

Defining the contrast independent component as

$$\boldsymbol{s} := \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{K} \sin\Big(-\frac{2\pi}{\lambda}(\alpha \boldsymbol{u} + \beta \boldsymbol{v})\Big), \tag{B.8}$$

Eq. (3.1), gives the approximate whitened kernel phase $\boldsymbol{x}$ for the observed object

$$\boldsymbol{x} \approx c^{-1}.\boldsymbol{s}. \tag{B.9}$$

The value of the MLE for $\boldsymbol{s}$, $\hat{\boldsymbol{s}} := \boldsymbol{s}(\hat{\alpha}, \hat{\beta})$ can be found by a grid search over α and β. From there, finding the MLE $\hat{c}^{-1}$ is equivalent to finding the value for c^{-1} that minimises the quantity

$$\begin{aligned}(\hat{\boldsymbol{x}} - \boldsymbol{y})^T(\hat{\boldsymbol{x}} - \boldsymbol{y}) &\approx (c^{-1}\hat{\boldsymbol{s}} - \boldsymbol{y})^T(\widehat{c^{-1}}\hat{\boldsymbol{s}} - \boldsymbol{y}) \\ &\approx 2(\widehat{c^{-1}})^2 \boldsymbol{s}^T\boldsymbol{s} - 2\widehat{c^{-1}}\boldsymbol{s}^T\boldsymbol{y} + \boldsymbol{y}^T\boldsymbol{y}.\end{aligned} \tag{B.10}$$

The MLE on the contrast can then be found by finding the value of $\widehat{c^{-1}}$ that brings the derivative over $\widehat{c^{-1}}$ of this last quantity to zero, and therefore by solving

$$\frac{d}{d(\widehat{c^{-1}})}(\widehat{c^{-1}})^2 \hat{\boldsymbol{s}}^T\boldsymbol{s} - 2\widehat{c^{-1}}\boldsymbol{s}^T\boldsymbol{y} + \boldsymbol{y}^T\boldsymbol{y} = 0. \tag{B.11}$$

Since this derivative is

$$\frac{d}{d(\widehat{c^{-1}})} 2(\widehat{c^{-1}})\hat{\boldsymbol{s}}^T\hat{\boldsymbol{s}} - 2\hat{\boldsymbol{s}}^T\boldsymbol{y}, \tag{B.12}$$

The MLE for the inverse contrast is

$$\widehat{c^{-1}} \approx \frac{\hat{\boldsymbol{s}}^T\hat{\boldsymbol{y}}}{\hat{\boldsymbol{s}}^T\hat{\boldsymbol{s}}}, \tag{B.13}$$

and so,

$$\hat{c} \approx \frac{\hat{\boldsymbol{s}}^T\hat{\boldsymbol{s}}}{\hat{\boldsymbol{s}}^T\boldsymbol{y}}. \tag{B.14}$$

The values of $\boldsymbol{s}$ can be precomputed, allowing for a quick grid search over α and β, and the analytical expression for the MLE over the contrast allows for quickly determining an approximate MLE $\widehat{c^{-1}}\hat{c} \approx \hat{\boldsymbol{x}}$. The results can also be subsequently refined by using the MLE from the grid search as the starting point of a gradient descent. It is however still too slow to constrain P_{FA} down to a very low level, say 10^{-6}, due to the time necessary to obtain the few million realisations necessary to reach such high levels of confidence, unless one has access to many processor cores. Determining a single MLE, using a combined AGS and gradient descent method takes about one second on a current generation Intel Xeon CPU (on a single core), with a discrete pupil model that yields around 1,000 kernel phases. Such a CPU is capable of simultaneously treating 48 threads, 10^6 realisations can thus be treated in under 60 hours.

The AGS presents a notable advantage compared to a classical grid search, since it only requires a search on a $N_\alpha \times N_\beta$ sized grid, rather than on a $N_\alpha \times N_\beta \times N_c$ one, greatly reducing the number of estimations that need to be made to obtain the MLE. It still requires more steps than a gradient descent with either a pre-set or a random starting point.

With this hybrid method, a grid search method becomes less computationally expensive. The AGS can therefore serve to determine how robust to local maximums other MLE determination methods are. This of special importance to the determination of the distribution of the test statistic of T_B under $\mathcal{H}_0$: to guarantee a given P_{FA}, it is necessary to determine the test statistic of T_B for more than P_{FA}^{-1} realisations. Thus, being able to examine the behaviour of faster methods that are more sensitive to local maximums can be very useful. In this case, I compared the distribution of T_B under $\mathcal{H}_0$ using MLEs determined with the AGS, to the distribution estimated using a gradient descent with a random starting point, to estimate the bias produced by the use of a fast, albeit imperfect MLE determination method.

This bias on the estimated distribution under $\mathcal{H}_0$ only goes in one direction: it can only lead to an underestimation of P_{FA}, making detection seem more certain than they really are. This is because, by definition of the MLE, any signature $\boldsymbol{z} \neq \hat{\boldsymbol{x}}$ that is used to compute the test statistic $T_B(\boldsymbol{y})$ yields a likelihood $\ell(\boldsymbol{z}; \boldsymbol{y}$ that is strictly lower than $\ell(\hat{\boldsymbol{x}}; \boldsymbol{y})$. Thus, a bad estimation of

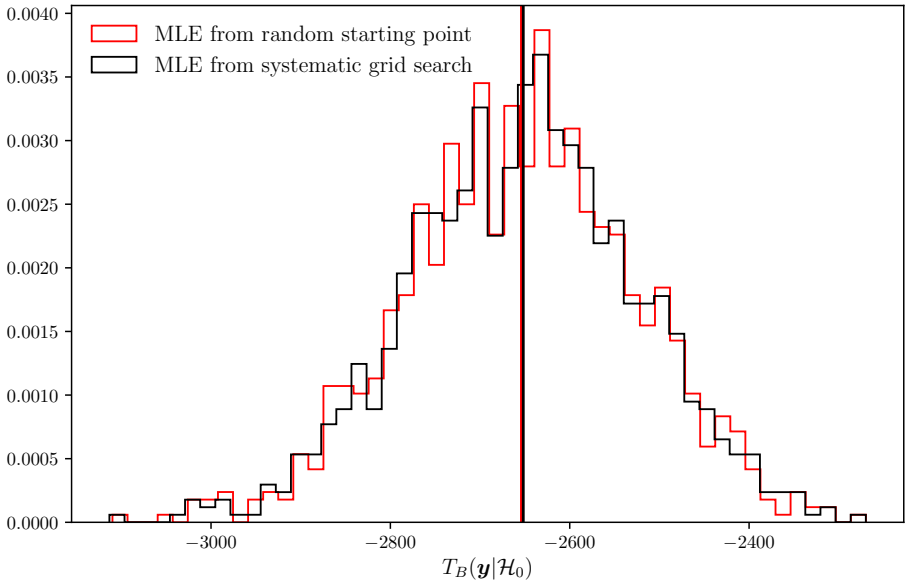

FIGURE B.1: The distribution of the test statistic of T_B under the null hypothesis $\mathcal{H}_0$ for a MLE $\hat{\boldsymbol{x}}$ estimated by a gradient descent, either with a starting point set randomly (red), or with a starting point obtained by a grid search (black). The values of $\boldsymbol{y}$ are the same random sample for both distribution, and the vertical bars represent the mean value of either distribution. The two methods yield distributions close enough for the random starting point not to be considered an issue.

the MLE will lead to an underestimation of the test statistic $T_B(\boldsymbol{y})$. Therefore, if the errors on the MLE are significant, the estimated distribution of $T_B(\boldsymbol{y})$ is biased under $\mathcal{H}_0$, leading to an underestimation of P_{FA}. It is therefore important to ensure than the MLE determination method used in each case does not produce a significant bias on the distribution of $T_B(\boldsymbol{y})$.

Fig. B.1 shows the distributions of the test statistic of T_B under $\mathcal{H}_0$ obtained by the AGS, and gradient descent with a random starting point (GDSP). The distribution obtained using the GDSP exhibits the slightest skew, but not enough to create an important error on the determination of P_{FA}. Thus, the fast GDSP is a valid method to estimate the distribution of the test statistic of T_B under $\mathcal{H}_0$.

Appendix C

Fundamental resolution limit: linear decomposition of a binary's signal and aberration degeneracy.

Around a given value, a function can be approximated as the sum of the first terms of its Taylor expansion. This decomposition of the signal of a binary as a series of Taylor polynomials was spurred by the possibility of developing detection methods that use the linear structure of a signal to detect it using a subspace detector proposed by Scharf and Friedlander (1994) which employs a GLR for which the MLE is obtained by a projection into the subspace spanned by the expected signature. The fact that the approximation breaks down at larger separations, and the relatively high number of dimensions of the basis that described the subspace spanned by the Taylor expansion of the phase of the visibility of a binary led us to employ an MLE obtained by a straightforward likelihood maximisation, which presents the advantage of offering an exact expression for this phase at every separation low enough for the signature to be captured by the discrete pupil model.

The case that concerns us here is the phase of the complex visibility of a binary. The Taylor coefficients are given by successive derivations of the phase of the complex visibility of a binary in Eq. (3.44), and the development up to the third order gives:

$$\begin{aligned}\angle \boldsymbol{\mathcal{V}}_0(u,v) = &-u\frac{c\alpha}{1+\frac{1}{c}} - v\frac{c\beta}{1+\frac{1}{c}} + \frac{u^3}{6c}\alpha^3\Big(1-\frac{6c^2}{(c+1)}\Big) + \frac{u^2v}{2c}\alpha^2\beta\Big(1-\frac{6c^2}{(c+1)}\Big)\\ &+ \frac{uv^2}{2c}\alpha\beta^2\Big(1-\frac{6c^2}{(c+1)}\Big) + \frac{v^3}{6c}\beta^3\Big(1-\frac{6c^2}{(c+1)}\Big).\end{aligned} \tag{C.1}$$

A linear basis can be associated with each of the term of the Taylor series. Since the phase signal is odd, only the odd order contribute to it. The basis spanned by the polynomials up to the fists order is

$$[\boldsymbol{u}, \boldsymbol{v}], \tag{C.2}$$

up to the third order, it is

$$[\boldsymbol{u}, \boldsymbol{v}, [\boldsymbol{u}^2\boldsymbol{v}, \boldsymbol{v}^2\boldsymbol{u}, \boldsymbol{u}^3, \boldsymbol{v}^3]], \tag{C.3}$$

with the power notation $\boldsymbol{a}^n$ denoting a vector that contain all of the terms of $\boldsymbol{a}$ elevated to the power n, and the notation $\boldsymbol{ab}$ the vector containing the element-wise product of $\boldsymbol{a}$ and $\boldsymbol{b}$. The terms up to the fifth order live in the basis

$$[\boldsymbol{u}, \boldsymbol{v}, [\boldsymbol{u}^2\boldsymbol{v}, \boldsymbol{v}^2\boldsymbol{u}, \boldsymbol{u}^3, \boldsymbol{v}^3], \boldsymbol{u}\boldsymbol{v}^4, \boldsymbol{u}^2\boldsymbol{v}^3, \boldsymbol{u}^3\boldsymbol{v}^2, \boldsymbol{u}^4\boldsymbol{v}, \boldsymbol{u}^5, \boldsymbol{v}^5]. \tag{C.4}$$

The two first-order terms are of particular interest: in the Fourier plane, a phase linearly dependent with the u and v coordinates is degenerate with a tip-tilt on the phase screen, which creates a shift on the image on the focal plane. Since tip-tilt in the Fourier plane is the consequence of a phase screen in the pupil plane, it is killed off by a kernels matrix.

At small separations, most of the signal of the binary is contained withing the tip-tilt term, i.e the Fourier phase can be approximated as a linear function of the u and v coordinates. The kernel matrix removes this linear components, which it "sees" as a tip-tilt aberration. Thus, the energy (or the squared norm of the vector) in the ker-phase of a binary signatures drops extremely fast as the separation decreases. For the JWST pupil discretised with the model shown in Fig. 3.5, at $0.4\lambda/D$, only 1% of the energy of the signal at 1 λ/D is left. This drops off to 0.05% at $0.2\lambda/D$, and 10^{-6} at $0.1\lambda/D$. If one equates the energy (the squared norm of the vector) of the ker-phase signal with the ease of detection, this means that a signal corresponding to a companion with contrast of 10 at a separation of $0.1\lambda/D$ is as difficult to detect as a companion of contrast 10^{-7} at a separation of $1\lambda/D$.

This is a fundamental limit, not only of kernel phase, but of any phase-based aberration robust detection methods at very small separations: when the sampling of the signature of interest is poor enough, the accessible signal (which is not tip-tilt degenerate) becomes extremely low. This meshes with experience regarding the resolution attainable using the phase of the Fourier visibility, be it by nulling, NRM or full pupil kernel-phase.[1]

[1] The amplitude of the signal of a binary follows a similar relation, and is therefore also sensitive to this effect.

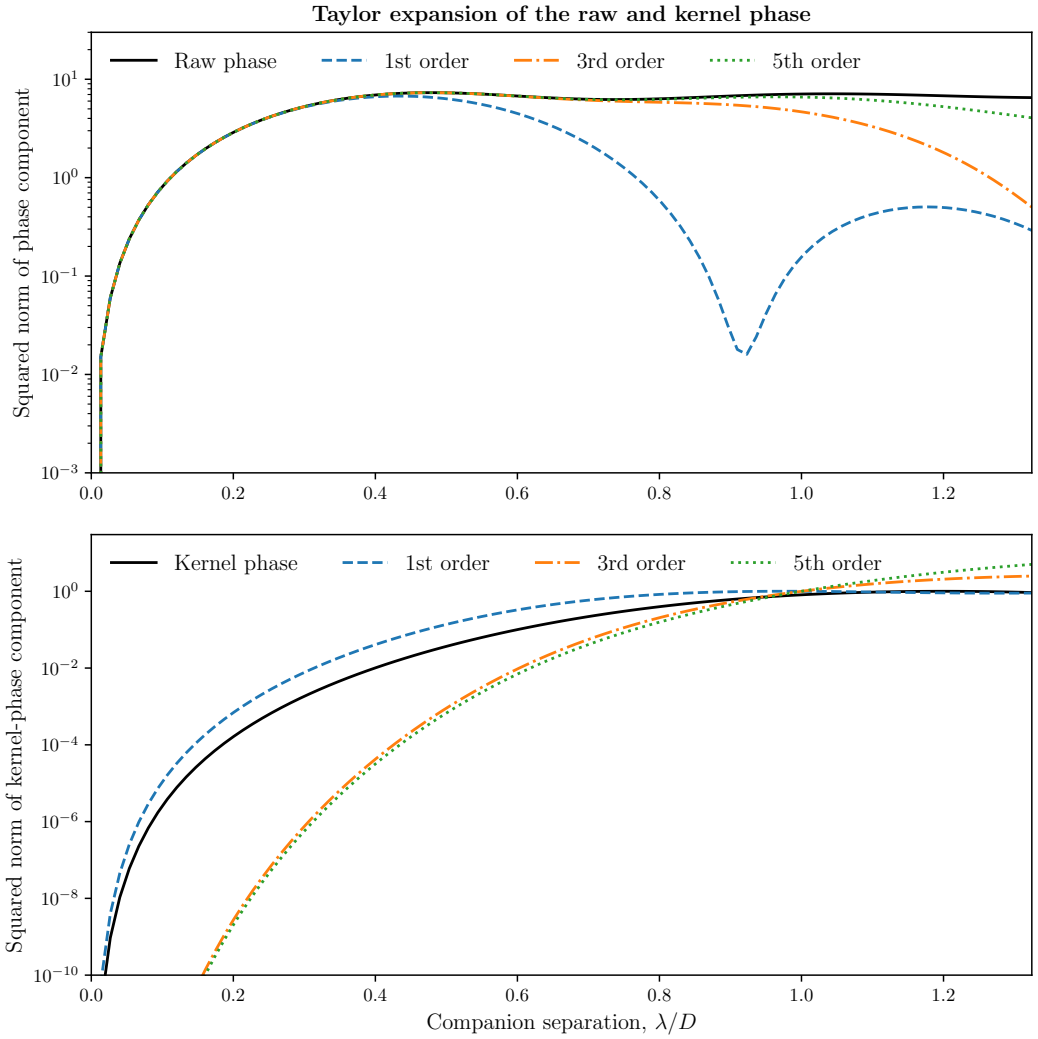

FIGURE C.1: Top panel: comparison of the squared norm of the phase of a binary and of the squared norm of its Taylor expansion to the 1st, 3rd and 5th orders, computed by projection of the bases from Eq. (C.2), Eq. (C.3) and Eq. (C.4). Bottom panel: kernel phase of the binary (solid black line), and projection of the phase *outside* of the subspace spanned by the aforementioned bases. The top panel indicates how well each basis can describe the phase signal, while the bottom panel compares the square norm of the phase signal projected either in the kernels space, or outside of the basis spanned by the approximation up to the first, third and fifth order. The bottom panel indicates that the loss in power of the kernel phase signal as the separation decreases is faster than the loss of power int he raw phase signal, and also that it matches closely the loss experienced by a signal with simplify the global tip-tilt component removed. this indicates that the bulk of the drop in detection performance at small separations is attributable to the loss of the tip-tilt signal.

A&A 630, A120 (2019)
https://doi.org/10.1051/0004-6361/201935247

Astronomy & Astrophysics

Kernel-phase detection limits

Hypothesis testing and the example of JWST NIRISS full-pupil images

A. Ceau[1], D. Mary[1], A. Greenbaum[2], F. Martinache[1], A. Sivaramakrishnan[3], R. Laugier[1], and M. N'Diaye[1]

[1] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France
e-mail: alban.ceau@oca.eu
[2] Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
[3] Space Telescope Science Institute, Baltimore, MD 21218, USA

Received 11 February 2019 / Accepted 7 August 2019

ABSTRACT

Context. The *James Webb* Space Telescope (JWST) will offer high angular resolution observing capability in the near-infrared with masking interferometry on the Near-Infrared Imager and Slitless Spectrograph (NIRISS), and coronagraphic imaging on the Near-Infrared Camera (NIRCam) and the Mid-Infrared Instrument (MIRI). Full-aperture kernel-phase-based interferometry complements these observing modes by allowing us to probe for companions at small angular resolution while preserving the telescope throughput.
Aims. Our goal is to derive both theoretical and operational contrast-detection limits for the kernel-phase analysis of JWST NIRISS full-pupil observations using tools from hypothesis testing theory. The study is immediately applied to observations of faint brown dwarfs with this instrument, but the tools and methods introduced here are applicable in a wide variety of contexts.
Methods. We construct a statistically independent set of observable quantities from a collection of aberration-robust kernel phases. Three detection tests based on these observable quantities are designed and analysed, all having the property of guaranteeing a constant false-alarm rate for phase aberrations smaller than about one radian. One of these tests, the likelihood ratio or Neyman-Pearson test, provides a theoretical performance bound for any detection test.
Results. The operational detection method considered here is shown to exhibit only marginal power loss with respect to the theoretical bound. In principle, for the test set to a false-alarm probability of 1%, companions at contrasts reaching 10^3 and separations of 200 mas around objects of magnitude 14.1 are detectable with a probability of 68%. For the brightest objects observable using the full pupil of JWST and NIRISS, contrasts of up to 10^4 at separations of 200 mas could ultimately be achieved, barring significant wavefront drift. We also provide a statistical analysis of the uncertainties affecting the contrasts and separations that are estimated for the detected companions.
Conclusions. The proposed detection method is close to the ultimate bound and offers guarantees on the probability of making a false detection for binaries, as well as on the error bars for the estimated parameters of the binaries that will be detected by JWST NIRISS. This method is not only applicable to JWST NIRISS but to any imaging system with adequate sampling.

Key words. instrumentation: high angular resolution – methods: data analysis – stars: low-mass – binaries: close – techniques: image processing – methods: statistical

1. Introduction

In the past few years, many nearby brown dwarfs have been discovered thanks to the Wide-field Infrared Survey Explorer (WISE) sky survey (Wright et al. 2010; Cushing et al. 2011; Schneider et al. 2015). These newly discovered objects present an observational challenge due to their intrinsically low luminosities. Some of them have been observed by the *Hubble* Space Telescope (HST), mostly for proper motion and parallax measurements (e.g. Marsh et al. 2013). While previous studies have searched for companions, they lacked the sensitivity in the optical and the near infrared to achieve high enough contrasts to detect very low-mass companions (e.g. Fontanive et al. 2018). High angular resolution observations are also possible from the ground using either adaptive optics or optical interferometry. Cool dwarfs are however intrinsically faint objects and therefore fall short of the requirements of either technique, unless assisted by laser guide stars (Bernat et al. 2010).

Issues limiting the quality of ground-based observations, such as sky background or atmospheric perturbations, can be alleviated by observing from space. When launched, the *James Webb* Space Telescope (JWST, Gardner et al. 2006) will be the largest ever space telescope, and will provide unparalleled sensitivity for studying faint, cool dwarfs. With a 6.5 m primary mirror, and an instrument suite covering the 0.6−25.5 μm wavelength range, the theoretical angular resolution of this telescope ranges from 20 to 800 mas. For a nearby object located less than 20 pc away, this translates to the ability to resolve structures present within a few astronomical units (AU) of the central source.

However, even for instruments capable of very high angular resolution, the glare from an object can drown out the light from faint surrounding structures. This issue is usually addressed by using coronagraphy and the instrumentation of the JWST offers several coronagraphs inside the Near-Infrared Camera (NIRCam) and the Mid-Infrared instrument (MIRI), with inner working angles ranging from 300 to 800 mas. To probe the innermost parts of nearby systems, inside the inner working angles of the coronagraphs, interferometry offers a viable alternative. In that scope, onboard JWST, the Near-Infrared Imager

and Slitless Spectrograph (NIRISS) offers the aperture masking interferometer (AMI) observing mode (Sivaramakrishnan et al. 2012) with a non-redundant mask (NRM) located in the instrument pupil wheel. The AMI enables the detection of objects with lower contrasts, but at narrower separations compared to what can be achieved by the JWST coronagraphs. The AMI is expected to have sufficient performance to address yet unanswered questions in the fields of active galaxy nuclei (AGNs; Ford et al. 2014), planetary formation, exoplanets (Artigau et al. 2014), and to facilitate follow-ups on astrometry measurements from the *Gaia* mission, or on ground-based extreme adaptive optics (AO) surveys. In the case of binary point sources in non-coronagraphic modes, contrast ratios as high as 10 mag (10^4) for the brightest companions at 130 mas can be attained using AMI (Sivaramakrishnan et al. 2012; Greenbaum et al. 2015, 2018).

The AMI achieves its best performance by taking advantage of self-calibrating observable quantities called closure phases (Jennison 1958). This technique, first developed for radio interferometry and later adapted to the optical regime (Baldwin et al. 1986) was adapted to single-dish telescopes using a non-redundant aperture mask. Initially used in seeing-limited observing conditions (Nakajima 1989), the technique eventually took advantage of the development of AO (Tuthill et al. 2006) allowing stabilised longer-exposure modes and providing the ability to observe fainter objects. Non-redundant mask interferometry is now routinely used and has led to a variety of studies (e.g. Sallum et al. 2015; Kraus et al. 2008, 2011).

Kernel phase generalises the idea of closure phase to apertures of arbitrary shapes, and can be reliably used when aberrations are smaller than about one radian (Martinache 2010). This method can therefore be used on images acquired using any instrument onboard JWST, provided that the instrument pupil geometry is accurately modelled. It is therefore useable on full-pupil images as well as on AMI/NRM closure phases. The Kernel method has already been used successfully to uncover new brown dwarf binaries with HST observations, as reported by Pope et al. (2013). Full-aperture kernel phase and AMI closure phase cover the same parameter space but with its lower throughput (~15%), AMI is suited to the observation of bright targets that would otherwise saturate the instrument, as well as to observations where aberrations are too important to fall into the linear regime covered by the kernel method.

Kernel- and closure phase rely on exploiting the phase of the Fourier transform (also referred to as the complex visibility) of the image. The image must satisfy the Nyquist-sampling requirement (platescale smaller than $0.5\,\lambda/D$), although small-grid dithering allows observers to reconstruct a Nyquist-sampled image for other filters. Saturation should be avoided, although recovery is still possible (Laugier et al. 2019). For a filter to be fully exploitable, its shortest wavelength must respect the sampling criterion. For the 6.5 m diameter of the primary mirror of JWST, this means that the filters compatible with a Kernel-phase analysis are:

- NIRCam in the short wavelength channel (0.6−2.3 μm), with a platescale of 31 mas pixel^{-1}: F212N
- NIRCam in the long wavelength channel (2.4−5.0 μm), with a platescale of 63 mas pixel^{-1}: F430M, F460M, F466N, F470N and F480M.
- NIRISS, with a platescale of 65 mas pixel^{-1} (STSCI 2018): F430M and F480M.
- MIRI, with a platescale of 110 mas pixel^{-1}: all filters but F770W and F780W.

Kernel detection limits for NIRCam have been computed by Sallum & Skemer (2019) for the F430M and F480M filters, as well a for NIRISS AMI in those same bands. The present work aims at setting a general statistical framework for the theoretical and operational detection limits of the Kernel approach, with focus on guarantees for the actual false-alarm rate of the implemented detection method. As for the practical results, we investigate various aspects of the detection limits achievable for full-aperture NIRISS observations in the F480M filter.

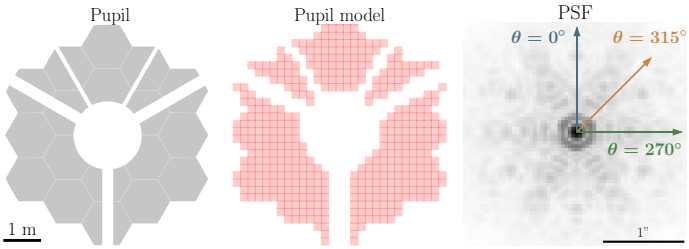

Fig. 1. *Left*: entrance pupil for JWST. *Centre*: discrete model of the pupil. The pupil is modelled by an array of subpupils, enabling the use of the kernel method. *Right*: simulated PSF for NIRISS using the 480M filter, represented using a non-linear colour scale. The coloured arrows represent the directions along which the simulated companions are placed.

Section 2 describes how kernel phases are constructed, presents the corresponding statistical model, and introduces three statistical tests that are later used to determine contrast detection limits. Section 3 shows how the method is applied to simulated images by JWST NIRISS. For several objects representative of the Y dwarfs discovered by WISE, this part highlights the need for estimating the noise covariance matrix, compares the performance of proposed detection tests, and analyses the statistical uncertainty resulting in the estimation of the parameters of the detected binaries. For the remainder of this paper, an italicised lowercase letter such as a denotes a real or complex number, a bold lowercase italicised letter such as $\boldsymbol{a}$ denotes a vector, a bold italicised uppercase letter such as $\mathbf{A}$ denotes a matrix, and a hat such as $\widehat{b}$ denotes the maximum likelihood estimate (MLE) of an unknown parameter b.

2. Kernel approach and statistical models

2.1. Kernel approach

The kernel framework introduced by Martinache (2010) describes diffraction-dominated images produced by the mostly continuous aperture of a telescope as if they were the interference pattern formed by a discrete array of virtual subapertures laid out on a regular grid of finite step. Although any pupil model can in principle yield kernel phases, using a regularly spaced grid allows the redundancy of the filled aperture to be encoded simply and effectively. The fidelity of the discrete representation of the continuous aperture increases with the density the grid. In practice however, the size of the grid step (s) translates into a cut-off frequency λ/s that is matched to the field of view over which the diffractive signal is recorded. The example of the discrete representation of the JWST entrance aperture along with an image of the theoretical point spread function (PSF) of the original aperture are shown in Fig. 1. The entrance pupil is a combination of the entrance pupil and of an additional pupil-plane mask, CLEARP.

Kernel phases are formed from a linear combination of the phase measured in the Fourier transform of the image. For a given wavelength, the discrete grid describing the original aperture also defines the sampling of the Fourier space via the

coordinates and redundancies of the different baselines. The values of the Fourier transform of an image for the selected spatial frequencies are collected in the complex visibility vector $\boldsymbol{v}$. The phase vector $\boldsymbol{\phi}$ is defined as the argument of the complex visibility,

$$\boldsymbol{\phi} := \angle \boldsymbol{v}. \tag{1}$$

In the optical path of a diffraction-limited instrument, unknown and potentially evolving aberrations result in a variable PSF that degrades the image quality. According to Martinache (2010), in the small aberration regime and for simple (i.e. non-coronagraphic) images, a linear model relates the phase $\boldsymbol{\phi}$ measured in the Fourier space to the true phase of the observed object $\boldsymbol{\phi_0}$ and to the aberration phase $\boldsymbol{\varphi}$ present across the aperture:

$$\boldsymbol{\phi} = \boldsymbol{\phi_0} + \mathbf{A}\boldsymbol{\varphi}, \tag{2}$$

where $\mathbf{A}$ is a phase transfer matrix, encoding how the aberrations in each subaperture will propagate to the Fourier phase of the image. Its properties depend on the discrete representation of the aperture. The discrete model of the aperture of JWST featured in Fig. 1 is made of $m = 452$ virtual subapertures, placed on a grid with a step size of 20 cm that form $n = 1363$ distinct baselines, resulting in a full rank phase transfer matrix $\mathbf{A}$ of dimensions 1363×452. The kernel matrix $\mathbf{K}$ is defined as a $p \times n$ matrix that verifies

$$\mathbf{K}\mathbf{A} = \mathbf{0}. \tag{3}$$

The kernel matrix cancels phase perturbation to the first order (Ireland 2013). With the chosen model, this matrix makes it possible to form a vector of kernel phases $\boldsymbol{k}$ of size $(p \times 1)$, with $p = 887$, defined as:

$$\boldsymbol{k} := \mathbf{K}\boldsymbol{\phi}. \tag{4}$$

The kernel matrix $\mathbf{K}$ represents the left-nullspace of the transfer matrix $\mathbf{A}$, and is computed from its singular value decomposition. The discrete representation of the aperture, the associated phase transfer matrix $\mathbf{A}$, and the kernel matrix $\mathbf{K}$ can be generated using a specially designed Python package called `XARA`[1], which also offers the basic tools to extract kernel phases from images.

2.2. Statistical modelling and hypothesis tests

Given a data image, how likely is it that a companion is present? The present study proposes to tackle this question through statistical hypothesis testing. A hypothesis test compares a test statistic (noted T) to a threshold (ξ), and has the general form

$$T(\boldsymbol{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi, \tag{5}$$

where $\boldsymbol{y}$ is the data under test (obtained from the image) and the test statistic $T(\boldsymbol{y})$ is a real random variable. In (5), the null hypothesis $\mathcal{H}_0$ (noise only) is accepted if $T(\boldsymbol{y}) < \xi$ and the alternative hypothesis $\mathcal{H}_1$ (noise + companion) is accepted otherwise. If the distribution of T can be known, the probability of false alarm can be controlled by the value of the test threshold ξ.

The performance of a detection test is given by its probability of false alarm (P_{FA}, the probability that a detection occurs under $\mathcal{H}_0$) and its probability of detection (P_{DET}, the probability that a detection occurs under $\mathcal{H}_1$):

$$\begin{aligned} P_{\mathrm{FA}} &:= \Pr(T(y) > \xi \,;\, \mathcal{H}_0), \\ P_{\mathrm{DET}} &:= \Pr(T(y) > \xi \,;\, \mathcal{H}_1). \end{aligned} \tag{6}$$

The *power* of a test is its P_{DET} at a given P_{FA}: the higher the P_{DET} for a given P_{FA}, the more powerful the test. It can be conveniently represented as a receiver operating characteristic (ROC) curve, P_{DET} as a function of P_{FA}.

Turning back to our detection problem, in the absence of noise the kernel phases can take the values

$$\begin{cases} \boldsymbol{k} = \mathbf{0}, & \text{if the target is centrosymmetric or} \\ \boldsymbol{k} = \mathbf{K}\boldsymbol{\phi_0}, & \text{if the target presents asymmetries.} \end{cases} \tag{7}$$

The noises affecting the images propagate into the Fourier phases and consequently into the kernel phases. As we see in the following section, the noise on the kernels can be modelled by a correlated Gaussian distribution with a covariance denoted $\boldsymbol{\Sigma}$. If this matrix is known, we can construct a vector $\boldsymbol{y}$ of "whitened" kernel phases which are decorrelated (hence independent), and similarly a vector $\boldsymbol{x}$ of whitened theoretical kernel phases corresponding to the signature of the target:

$$\boldsymbol{y} := \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{k}, \tag{8}$$

$$\boldsymbol{x} := \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{K}\boldsymbol{\phi_0}. \tag{9}$$

This leads to the following statistical hypotheses:

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y} = \boldsymbol{\epsilon} \\ \mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon} \end{cases}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{10}$$

where $\boldsymbol{\epsilon}$ is a $p \times 1$ noise vector with independent and identically distributed Gaussian entries (thanks to the whitening), and $\mathcal{N}(\mathbf{0}, \mathbf{I})$ denotes the standard normal distribution (the covariance of $\boldsymbol{\epsilon}$ is the identity matrix, $\mathbf{I}$).

2.2.1. Known signature in white Gaussian noise

For the problem defined in (10), the most powerful test is the likelihood ratio (LR), or Neyman-Pearson (NP) test (Neyman & Pearson 1933). For this test, the companion signature $\boldsymbol{x}$ must be known. The NP test is defined as

$$\frac{\ell(\boldsymbol{x};\boldsymbol{y})}{\ell(\mathbf{0};\boldsymbol{y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{11}$$

where $\ell(\boldsymbol{x};\boldsymbol{y})$ is the likelihood of the signature $\boldsymbol{x}$ given the data $\boldsymbol{y}$ and η is an adjustable threshold. For the Gaussian white noise considered here, the likelihood is (Scharf & Friedlander 1994)

$$\ell(\boldsymbol{x};\boldsymbol{y}) = (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^T(\boldsymbol{x}-\boldsymbol{y})\right), \tag{12}$$

with p being the length of the kernel phase vector. The likelihood under $\mathcal{H}_0$ can be obtained from Eq. (12) by taking $\boldsymbol{x} = \mathbf{0}$. Combining Eqs. (12) and (11) under $\mathcal{H}_0$ and $\mathcal{H}_1$ gives the test

$$\exp\left(-\frac{1}{2}(\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{y}^T\boldsymbol{x})\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{13}$$

Taking the logarithm of Eq. (13) and noting $\xi := \eta + \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}$ leads to the test

$$T_{\mathrm{NP}}(\boldsymbol{y},\boldsymbol{x}) = \boldsymbol{y}^T\boldsymbol{x} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{14}$$

[1] XARA is available at http://github.com/fmartinache/xara

Hence, the NP test amounts to comparing the dot product of the data with the signature to a threshold. The distribution of $T_{\rm NP}$ can be analytically determined under $\mathcal{H}_0$ and $\mathcal{H}_1$:

$$\begin{cases} \mathcal{H}_0 : T_{\rm NP}(\boldsymbol{y}) \sim \mathcal{N}(0, \boldsymbol{x}^T\boldsymbol{x}), \\ \mathcal{H}_1 : T_{\rm NP}(\boldsymbol{y}, \boldsymbol{x}) \sim \mathcal{N}(\boldsymbol{x}^T\boldsymbol{x}, \boldsymbol{x}^T\boldsymbol{x}). \end{cases} \tag{15}$$

Using $\mathcal{N}(0,1)$ to denote a standard normal variable and $\mathcal{F}_{\mathcal{N}}$ its cumulative distribution function (CDF), $P_{\rm FA}$ and $P_{\rm DET}$ for $T_{\rm NP}$ can be derived using their definitions in Eq. (6) and the distributions in Eq. (15)

$$\begin{cases} P_{\rm FA}^{T_{\rm NP}}(\xi) = 1 - \mathcal{F}_{\mathcal{N}}\left(\dfrac{\xi}{\sqrt{\boldsymbol{x}^T\boldsymbol{x}}}\right), \\ P_{\rm DET}^{T_{\rm NP}}(\xi) = 1 - \mathcal{F}_{\mathcal{N}}\left(\dfrac{\xi - \boldsymbol{x}^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^T\boldsymbol{x}}}\right), \end{cases} \tag{16}$$

which, for the purpose of plotting ROC curves, combine to

$$P_{\rm DET}^{T_{\rm NP}}\left(P_{\rm FA}^{T_{\rm NP}}\right) = 1 - \mathcal{F}_{\mathcal{N}}\left(\mathcal{F}_{\mathcal{N}}^{-1}(1 - P_{\rm FA}^{T_{\rm NP}}) - \sqrt{\boldsymbol{x}^T\boldsymbol{x}}\right). \tag{17}$$

This test is the most powerful for the considered model, and serves as the benchmark against which any other detection test can be evaluated.

Implementing the NP test (Eq. (14)) requires knowledge of the target signature $\boldsymbol{x}$ (namely, contrast and position if $\boldsymbol{x}$ corresponds to a companion). In practical situations however, $\boldsymbol{x}$ is often partially or even fully unknown. This leads us to consider the statistical model

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y} = \boldsymbol{\epsilon}, \\ \mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon}, \quad \boldsymbol{x} \in \mathcal{X}\,, \end{cases} \tag{18}$$

where $\mathcal{X}$ is a space describing some prior information about $\boldsymbol{x}$. Below, we consider two cases: a completely unknown signature ($\mathcal{X} = \mathbb{R}^p$) and the signature of a binary with unknown contrast and separation ($\mathcal{X}$ is then the space spanned by all possible binary signatures). A classical approach when some parameters describing the target $\boldsymbol{x}$ are unknown is to inject its MLE (denoted $\widehat{\boldsymbol{x}}$) in place of $\boldsymbol{x}$ in the LR of Eq. (11). The MLE is defined by

$$\widehat{\boldsymbol{x}} := \underset{z \in \mathcal{X}}{\operatorname{argmax}}\ \ell(z; \boldsymbol{y}), \tag{19}$$

and injecting the MLE in the LR leads to the so-called generalised likelihood ratio (GLR) defined as

$$\frac{\max\limits_{z\in\mathcal{X}} \ell(z;\boldsymbol{y})}{\ell(\boldsymbol{0};\boldsymbol{y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta \quad \Leftrightarrow \quad \frac{\ell(\widehat{\boldsymbol{x}};\boldsymbol{y})}{\ell(\boldsymbol{0};\boldsymbol{y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{20}$$

2.2.2. Completely unknown $\boldsymbol{x}$ signature

If we assume as a worst-case situation that nothing is known about the signature $\boldsymbol{x}$, we have $\mathcal{X} = \mathbb{R}^p$. The likelihood in (12) is maximised for $\widehat{\boldsymbol{x}} = \boldsymbol{y}$, and injecting this value in Eq. (20) yields

$$\frac{\exp\left(-\dfrac{1}{2}(\boldsymbol{y}-\boldsymbol{y})^T(\boldsymbol{y}-\boldsymbol{y})\right)}{\exp\left(-\dfrac{1}{2}(\boldsymbol{y}-\boldsymbol{0})^T(\boldsymbol{y}-\boldsymbol{0})\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi', \tag{21}$$

with ξ' being a threshold. Taking the logarithm of Eq. (21), we obtain the test:

$$T_{\rm E}(\boldsymbol{y}) := \|\boldsymbol{y}\|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \tag{22}$$

This test uses the measured squared norm of the signal as a test statistic and is called an energy detector (hence $T_{\rm E}$). Its statistic is distributed as:

$$\begin{cases} \mathcal{H}_0 : T_{\rm E}(\boldsymbol{y}) \sim \chi^2_p(\lambda^2 = 0), \\ \mathcal{H}_1 : T_{\rm E}(\boldsymbol{y}) \sim \chi^2_p(\lambda^2 = \boldsymbol{x}^T\boldsymbol{x}). \end{cases} \tag{23}$$

Using $\mathcal{F}_{\chi^2_p(\lambda^2)}$ to denote the CDF of a $\chi^2_p(\lambda^2)$ random variable with p degrees of freedom and non-centrality parameter λ, we obtain

$$\begin{cases} P_{\rm FA}^{T_{\rm E}}(\xi) = 1 - \mathcal{F}_{\chi^2_p(0)}(\xi), \\ P_{\rm DET}^{T_{\rm E}}(\xi) = 1 - \mathcal{F}_{\chi^2_p(\boldsymbol{x}^T\boldsymbol{x})}(\xi). \end{cases} \tag{24}$$

We note that test $T_{\rm E}$ in Eq. (22) was previously used in the literature, for example by Zwieback et al. and Le Bouquin & Absil (2016; 2012) (although not identified as a GLR), with the $P_{\rm FA}$ reported in Eq. (24).

The expressions above combine into

$$P_{\rm DET}^{T_{\rm E}}\left(P_{\rm FA}^{T_{\rm E}}\right) = 1 - \mathcal{F}_{\chi^2_p(\boldsymbol{x}^T\boldsymbol{x})}\left(\mathcal{F}^{-1}_{\chi^2_p(\boldsymbol{0})}\left(1 - P_{\rm FA}^{T_{\rm E}}\right)\right). \tag{25}$$

Indeed, this test does not exploit any prior knowledge of the structure of the object to be detected and can thus be seen as providing a lower bound for the detection performance.

2.2.3. Signature of a binary

Repeated observations of gravitationally interacting multiple systems is the only means by which unambiguous dynamical masses can be determined. Because they make it possible to resolve asymmetries near or even slightly below the diffraction limit, which translates into small orbital distances, NRM closure or kernel phase (Kraus et al. 2008; Huélamo et al. 2011; Lacour et al. 2011) and full-aperture kernel phase (Pope et al. 2013; Laugier et al. 2019) are particularly suited to the observation of unequal-brightness, low-mass binary systems.

At any instant, a binary system is characterised by three parameters: the angular separation ρ of the companion relative to the primary, its position angle θ, and a contrast c, defined here as the luminosity ratio of the primary over the secondary. Our simulations assume that the position angle is measured in the image relative to the axis pointing up (represented by a blue arrow in the right hand panel of Fig. 1), and increases counterclockwise. Actual observations also have to take into account the orientation of the telescope to project the apparent position angle onto the celestial sphere to combine observations at multiple epochs.

As an intermediate step, it is also convenient to use a Cartesian coordinate system in which the location of the secondary is given by (α, β). If the binary system is made of two individually unresolved point sources, its intensity distribution O can be modelled as a pair of Dirac distributions:

$$O(x, y) = \delta(x, y) + c^{-1}\delta(x - \alpha, y - \beta). \tag{26}$$

The complex visibility $\boldsymbol{v}$ associated to this object is the 2D Fourier transform of Eq. (26) (van Cittert 1934; Zernike 1938), that is,

$$\boldsymbol{v}(\boldsymbol{u}, \boldsymbol{v}) = 1 + c^{-1}\exp\left(-i\frac{2\pi}{\lambda}(\alpha\boldsymbol{u} + \beta\boldsymbol{v})\right). \tag{27}$$

We reiterate that in the alternative hypothesis defined in Eq. (18), $\boldsymbol{x} = \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{K}\boldsymbol{\phi_0}$, where $\boldsymbol{\phi_0} = \angle\boldsymbol{v}$. This leads to the parametric hypothesis:

$$\mathcal{H}_1 : \boldsymbol{y} = \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{K}\angle\left(1 + c^{-1}\exp\left(-i\frac{2\pi}{\lambda}(\alpha\boldsymbol{u} + \beta\boldsymbol{v})\right)\right) + \boldsymbol{\epsilon}. \tag{28}$$

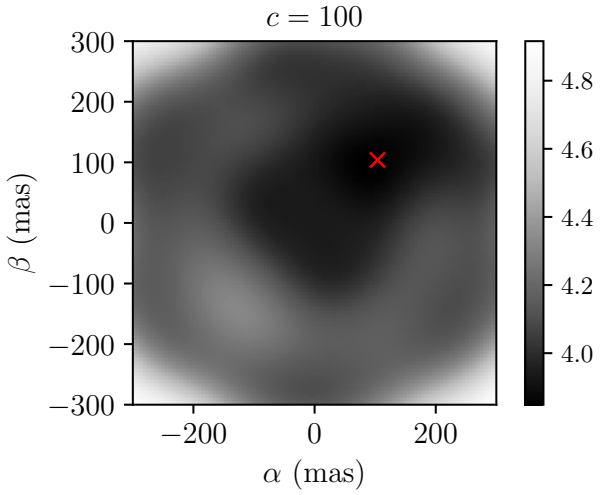

Fig. 2. Map of the likelihood that is maximised in Eq. (29) for a data vector $\boldsymbol{y}$ accounting for a realistic covariance matrix $\boldsymbol{\Sigma}$ for JWST NIRISS. The companion signature has parameters $\alpha = \beta = 104$ mas (red cross) and $c = 100$.

Under $\mathcal{H}_1$, there are three free parameters: α, β, and c, so the MLE is now

$$\widehat{\boldsymbol{x}} := \underset{\alpha,\beta,c}{\operatorname{argmax}}\ \ell(c,\alpha,\beta;\boldsymbol{y}),$$
$$= \underset{\alpha,\beta,c}{\operatorname{argmax}}\ e^{-\frac{1}{2}\left\| \boldsymbol{y} - \boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{K}\angle\left(1 + c^{-1}e^{-i\frac{2\pi}{\lambda}(\alpha\boldsymbol{u}+\beta\boldsymbol{v})}\right)\right\|^2}. \quad (29)$$

Finding the MLE is equivalent to minimising the argument of the exponential. This minimisation cannot be done analytically but numerical methods can be used to compute $\widehat{\boldsymbol{x}}$, as explained below. Injecting (29) in (20) gives the test

$$\frac{\exp\left(-\frac{1}{2}(\boldsymbol{y}-\widehat{\boldsymbol{x}})^T(\boldsymbol{y}-\widehat{\boldsymbol{x}})\right)}{\exp\left(-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{y}\right)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \quad (30)$$

equivalent to

$$T_{\mathrm{B}}(\boldsymbol{y}) := 2\boldsymbol{y}^T\widehat{\boldsymbol{x}} - \widehat{\boldsymbol{x}}^T\widehat{\boldsymbol{x}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \xi. \quad (31)$$

We note that this detection problem is similar to "case VII" of Scharf & Friedlander (1994), where the detection procedure also relies on the ML estimation of the signal of interest. In this latter study, however, the signature $\boldsymbol{x}$ is assumed to reside in a linear subspace (independent from the nuisance subspace), which is not the case here.

As mentioned above, the MLE $\widehat{\boldsymbol{x}}$ must be found numerically. Figure 2 illustrates, for one realisation of $\boldsymbol{\epsilon}$, an example of the value of the likelihood for a fixed contrast as a function of position angles α and β. It is apparent that the likelihood function is multimodal, so the minimisation strategy must be able to avoid local minima. A brute force search on a finely discretised grid of the parameter space is possible but comes at a large computation cost. Efficient numerical methods for solving multimodal problems exist, such as for instance the Monte Carlo Markov chains method with simulated annealing (Andrieu et al. 2003) or nested sampling (Skilling 2004).

Because the distribution of T_{B} involves the unknown distribution of the MLE estimate $\widehat{\boldsymbol{x}}$, it cannot be characterised analytically. However, as we see in the following section, this distribution can be estimated by Monte Carlo simulations, allowing us to accurately establish the relationship between the false-alarm probability $P_{\mathrm{FA}}^{T_{\mathrm{B}}}$ of this test and the threshold ξ in Eq. (31).

As an important final remark, we underline that the false-alarm probabilities of the considered tests are independent of the power of the phase perturbations $\boldsymbol{\varphi}$ (at least as long as the linear model in Eq. (2) holds, that is, for phase perturbation below $\approx$1 radian). This is clear from expressions (16) and (24) for tests T_{NP} and T_{E}; this is also the case for test T_{B} because the phase perturbation is cancelled by the operator $\mathbf{K}$ and does not affect the test statistic. This means that the false-alarm rate of these tests remains constant in case of fluctuating aberrations, which is a desirable feature in practice.

2.2.4. Likelihoods, likelihood ratios, and χ^2 intervals

The test statistic T_{B} can be interpreted in terms of χ^2-derived intervals as follows. Let $\hat{\boldsymbol{x}}$ be some model obtained by some fit on data $\boldsymbol{y}$. The χ^2 score corresponding to this fit is

$$T_{\chi^2}(\hat{\boldsymbol{x}},\boldsymbol{y}) := \sum_{k=1}^{N}(\hat{\boldsymbol{x}}_k - \boldsymbol{y}_k)^2 = (\hat{\boldsymbol{x}}-\boldsymbol{y})^T(\hat{\boldsymbol{x}}-\boldsymbol{y}). \quad (32)$$

Considering the likelihood in Eq. (12), this shows that if $\boldsymbol{y}$ is Gaussian with mean $\hat{\boldsymbol{x}}$, the score in Eq. (32) is indeed a χ^2_p random variable. Now, the test statistics T_{B} can be rewritten as

$$T_{\mathrm{B}} = 2\boldsymbol{y}^T\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}^T\hat{\boldsymbol{x}} = \boldsymbol{y}^T\boldsymbol{y} - \left((\hat{\boldsymbol{x}}-\boldsymbol{y})^T(\hat{\boldsymbol{x}}-\boldsymbol{y})\right) \quad (33)$$
$$= T_{\chi^2}(\mathbf{0},\boldsymbol{y}) - T_{\chi^2}(\hat{\boldsymbol{x}},\boldsymbol{y}), \quad (34)$$

which shows that T_{B} can be interpreted as the reduction in the sum of squared residuals when comparing the null hypothesis to the considered model.

For the sake of accurately controlling the false-alarm rate, we note however that T_{χ^2} in Eq. (32) may not be distributed as a χ^2_p variable because $\hat{\boldsymbol{x}}$ is a random variable. Actually, the true distribution of T_{χ^2} may not be known analytically, and a Monte Carlo procedure (such as that mentioned in Sect. 2.2.3 for the estimation of the correspondence between the P_{FA} and threshold for T_{B}) is required.

3. Results

The tests with the performance analyses presented in Sect. 2 are very general: considering a different aperture and instrumental noise simply amounts to replacing $\mathbf{A}$, $\mathbf{K}$, and $\boldsymbol{\Sigma}$ in the equations. We focus now on their specific application to JWST NIRISS full-pupil images (see Table 1).

3.1. Dataset and considered targets

We applied the three detection tests previously introduced to a series of simulated JWST/NIRISS datasets, replicating the observing scenario of archetypal ultracool Y-type brown dwarfs. While their multiplicity rate is currently unknown, more than 25 such objects have been discovered less than 20 pc away, mostly by the WISE mission (Kirkpatrick et al. 2011). At 20 pc, the theoretical angular resolution of JWST for $\lambda = 4.8\,\mu$m translates into an orbital distance of 3 AU: interferometric observations

Table 1. Detector and targets characteristics used to compute the covariance of the kernel phases extracted from our JWST NIRISS simulated dataset.

Read noise (e^-)	14.849
Flat field error	0.01%
Dark current (e^- s^{-1})	0.04
Total integration time (s)	2400
Number of frames	15
Gain (e^-/ADU)	1.00
Jitter value (mas)	7.0
Integration time	40 min
Number of photons (W2 mag = 15.4)	3.723×10^6
Number of photons (W2 mag = 14.1)	1.1181×10^7

will make it possible to probe within the first few AU of most known Y dwarfs.

JWST NIRISS images of Y dwarfs are simulated to evaluate the performance of the detection tests, using the `ami_sim`[2] package (Greenbaum et al. 2016), corresponding to a 40 min integration on target and a 40 min integration on a perfect calibrator. Frames are simulated in full-pupil mode, using the F480M filter, for two different "W2" magnitudes: 15.4 and 14.1. The W2 magnitude is the apparent magnitude in the band selected by the W2 ($\lambda = 4.6\,\mu$m) WISE filter (Wright et al. 2010). For these objects, companions are placed at a single position angle $\theta = 315°$ (materialised by the orange arrow in the PSF shown in Fig. 1). The simulated companions lie at separations of $\rho = 73$ mas ($\approx 0.5\lambda/D$ @ $\lambda = 4.8\,\mu$m) or $\rho = 147$ mas ($\approx \lambda/D$ @ $\lambda = 4.8\,\mu$m), and have contrasts $c = 10$, $c = 20$, $c = 50$, or $c = 100$, leading to a total of eight possible signatures.

For any given target, a calibration frame is simulated and we assume no calibration error (stable wavefront, calibrator with the same spectrum and brightness as the Y dwarf). To comply with a real situation, kernel phases are not extracted directly from the simulated image: the frames are recentred, cropped to a size of 64×64 pixels and apodized by a super-Gaussian mask (see Eq. (2) of Laugier et al. 2019) of 30 pixels in radius to weigh down the edges of the image.

3.2. Modelling the errors

Two types of errors affect kernel phases and the outcome of the statistical tests described in Sect. 2. First are statistical errors induced by random noises whose overall impact can be captured in the acquisition or the synthesis of a global covariance matrix. Second are systematic errors resulting from the imperfect modelling by the kernel framework of the broadband, long-exposure, and diffractive nature of images. The subtraction of kernel phases acquired on a point source theoretically accounts for this systematic error. However, in practice, wavefront drifts between observations will result in unaccounted-for residual errors referred to as systematic errors (Ireland 2013).

To estimate the potential impact of systematic errors induced by wavefront drift, we rely on Perrin et al. (2018) who predict that over a timescale of two hours, JWST drifts will result at most in a 16 nm rms wavefront across the entire pupil[3]. We used the ten OPD maps distributed with the `webbpsf` package, scaled down to correspond to the predicted rms to produce images resulting in ten distinct kernel-phase realisations. The dispersion of kernel phases across these realisations was used to estimate the magnitude of the calibration residual. In the bright target scenario (W2 mag = 14.1) introduced in Sect. 3.1, this calibration residual accounts for about 14 % of the total noise variance. As is shown further below, this systematic error has a small impact when observing faint targets.

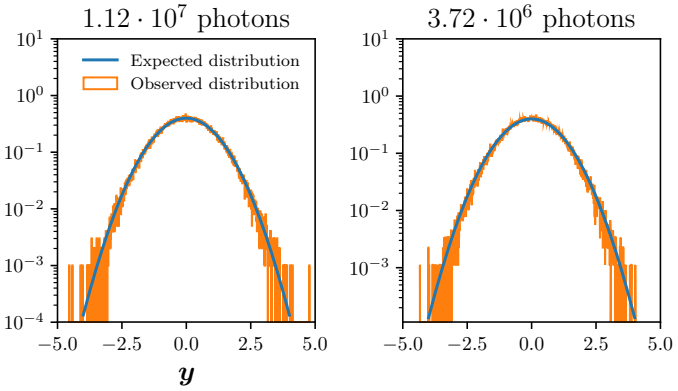

Fig. 3. Histogram of the values of the whitened kernel phases for the calibration images (orange). Standard, normal distribution (blue). *Left-panel*: higher flux regime. *Right panel*: lower flux regime. The distribution of whitened kernel phases obtained in practice is accurately described by the theoretical normal distribution considered in Eq. (10).

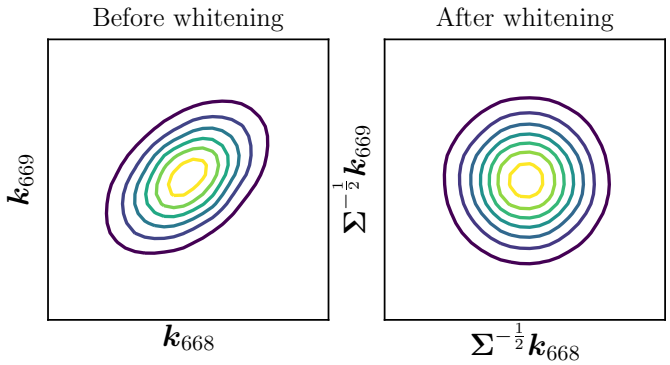

Fig. 4. Isocontours of the joint PDF of the 668th and 669th elements of the kernel-phase vector $\boldsymbol{k}$, $\boldsymbol{k}_{668}$ and $\boldsymbol{k}_{669}$, before (*left*) and after (*right*) whitening. The PDFs are estimated using 10^5 noise realisations.

3.3. Covariance estimation

Whereas simulated images used in the analysis include all the previously listed noises, experience has shown us that, apart from calibration residuals, the covariance matrix can accurately be estimated using the three dominant noises: photon, readout, and dark current. Figure 3 indeed shows that after whitening by this simpler covariance, the distribution of kernel-phases is indistinguishable from a normal distribution of standard deviation 1.

The effect of the whitening is further illustrated in Fig. 4, which shows how previously noise-correlated kernel phases (left panel) are indeed made statistically independent (right panel). The thus-whitened observables can indeed be reliably used as input for the different statistical tests introduced in Sect. 2.2.

In practice, the covariance $\boldsymbol{\Sigma}$ is estimated using Monte Carlo simulations. An accurate estimation requires a number of

[2] `ami_sim` is available at `https://github.com/agreenbaum/ami_sim`

[3] Perrin et al. (2018) predict that large variations in slew angle will result in the most important variations, as the primary mirror regains thermal equilibrium over the course of days.

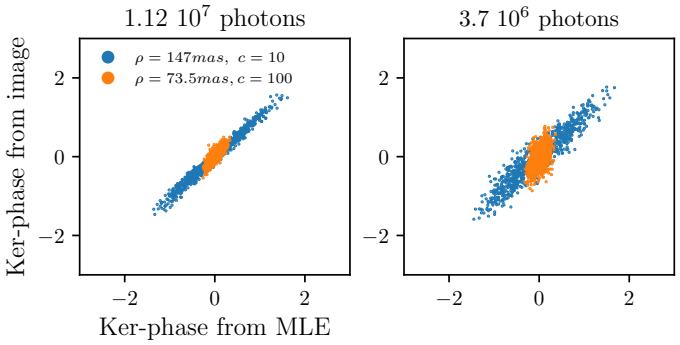

Fig. 5. Kernel phases of the recovered signature (x axis) against the true kernel phases of the injected binary (y axis). *Left panel*: high flux regime. *Right panel*: low flux regime. The worst S/N situation (ρ = 73.5 mas, c = 100) is in orange, and the best S/N (ρ = 147 mas, c = 10) is in blue.

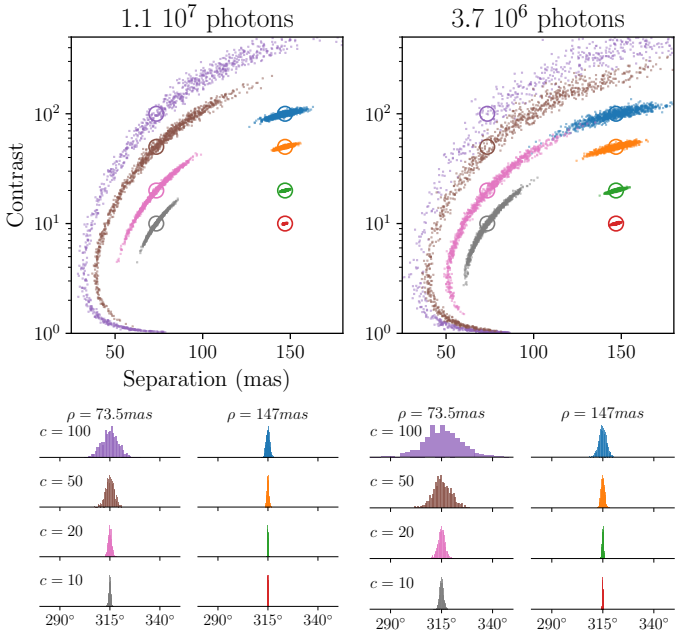

Fig. 6. Error on the recovered parameters. *Top panels*: circles represent the separation and contrasts of the injected signature, and each column represents one flux regime. *Bottom panels*: error on the estimated angle θ. The parameters ρ and c of each injected signature are represented as coloured circles on the top panels, while θ is fixed at 315° for all signatures. The same colour code is used for every panel, with each colour corresponding to an injected signature. Each dot on the top panel represents the parameters estimated for a single realisation of the noise.

simulated frames much greater than the total number of kernels; we used 10^5 frames for 887 kernels in our case.

Calibrated kernel phases are obtained by subtracting the kernel phases of a calibrator from those of the target in order to remove kernel model imperfections. Since the same flux is assumed for both observations, they share the same covariance. The covariance of the calibrated kernel-phase vector is therefore twice the covariance $\mathbf{\Sigma}_{\rm est}$ estimated from the MC simulations.

To account for unknown calibration errors reported in NRM-interferometry as well as in full aperture kernel phases that result in a kernel-phase bias, one commonly used solution has been to artificially inflate the experimental variance by adding an additional term whose overall magnitude is adjusted during the model fit (e.g. Martinache et al. 2009). The OPD maps introduced in Sect. 3.2 make it possible to estimate the magnitude of this bias *a priori*. Proper treatment of the calibration would require the subtraction of an estimate of the calibration term, using either the POISE algorithm of Ireland (2013) or the KL decomposition approach described by Kammerer et al. (2019) that relies on the observation of multiple calibration sources. Here we estimate the impact of an unaccounted-for calibration error on the contrast detection limits by adding the residual determined after analysis of the simulation that included the OPD maps to the diagonal of the covariance. To pursue the possibly covariated effects would require the computation of a distinct covariance matrix from a large number of distinct realisations of telescope drifts. For the faint brown dwarf case that motivates this study, the impact of the calibration error is small, and therefore we chose not to pursue the non-diagonal terms.

3.4. Parameter estimation

Detecting a companion using the operational binary test $T_{\rm B}$ requires the determination of the MLE $\widehat{\boldsymbol{x}}$. This requires estimation of the parameters ρ, θ, and c from the whitened kernel phases $\boldsymbol{y}$ in Eq. (28). The distribution of the parameters can be estimated by generating, for each considered signature, a large number of noisy kernel phases and estimating the parameters. In practice, a global optimisation algorithm can be used. For the purpose of making a large number of simulations, we assume that the algorithm has localised the region in which the global minimum is situated (the darkest region in Fig. 2). In this setting, the minimum can be found by a gradient descent algorithm.

In the following, we use the algorithm described by Branch et al. (1999), as implemented in `scipy.optimize.leastsquares`, which uses the local gradient and optimises for the direction descent and step size. The initialisation of the algorithm corresponds to the parameters of the injected companion. This method is suited for the determination of contrast limits thanks to its speed. We checked that we obtained very similar results with a (computationally more expensive) systematic grid search that would typically be used in practice[4].

Figure 5 shows the recovered kernel phases as a function of the kernel phases of simulated images for different separation, contrast, and flux regimes. The fit remains relatively consistent for each case, with scatter becoming predictably more important as the S/N decreases (the S/N is affected by the contrast, the separation, and the total flux in the image).

All of the signatures presented in Fig. 6 are detectable by the $T_{\rm B}$ with $P_{\rm FA} < 10^{-3}$. The shape of the 2D distribution of the estimated separation ρ and contrast c reproduces what was for instance reported by Pravdo et al. (2006) in the context of NRM observations: at angular separations smaller than λ/D, estimates for the contrast and the angular separation are strongly correlated.

Figure 6 also shows that two regimes can be distinguished. For a companion at $\rho \approx \lambda/D$ (for JWST λ/D = 152 mas @ λ4.80 μm), all parameters are well constrained, while for a companion at $\rho < \lambda/D$, the contrast and the angular separation cannot be well constrained simultaneously. In practice, this means that the estimation of the position of a companion using kernel-phases when the expected angular separation is smaller than λ/D can be further constrained by an independent measurement of the

[4] The gradient descent procedure is indeed only applicable in the context of the determination of detection limits by a Monte Carlo method.

luminosity of the companion at a different epoch, when $\rho > \lambda/D$. This property can be particularly useful in the case of objects with high eccentricities or inclinations.

A study of the consequences of the uncertainties on parameters and correlations on the orbit that can be fitted using the Kernel method on NIRISS images is out of the scope of this paper; this should be the object of future work, along with recommendations of optimum observing strategies in regards to the uncertainties on measured orbital parameters.

3.5. Detection and contrast performance

Firstly, we validate the theoretical relations predicting the performance of the NP test T_{NP} (Eq. (17)) and of the energy detector T_{E} (Eq. (24)), and determine the actual performance of T_{B} (Eq. (31)). For that purpose, we perform Monte Carlo simulations consisting of 2000 realisations[5] of $\boldsymbol{y}$ under $\mathcal{H}_0$ and under $\mathcal{H}_1$ for a given signature $\boldsymbol{x}$ (cf. Eq. (27)).

All of the detection limits are shown for $P_{\mathrm{FA}} = 1\%$ and $P_{\mathrm{DET}} = 68\%$. In terms more frequently encountered in astronomy publications, this is equivalent to having a 68% chance of making a $\approx 2.3\sigma$ detection.

On each realisation, we perform each of the three tests using the kernels operator $\mathbf{K}$ and the covariance matrix $\mathbf{\Sigma}$ estimated as in Sect. 3.3.

Figure 7 presents our results in the form of ROC curves, which provide a graphical representation of the power of each test. It can be seen that the dashed curves representing the theoretical ROCs accurately match the solid lines corresponding to the performance achieved in practice. As expected, T_{NP} appears to be the most powerful of the three tests (this test corresponds to the upper performance bound) and T_{E} the least powerful of the three (this test uses no prior information on the target signature and can be seen as a lower bound). The performance of T_{B} logically lies in between, but much closer to the upper than to the lower bound.

The detection limits for the three tests T_{NP}, T_{E}, and T_{B} are represented in Fig. 8 across a range of contrasts and separations, for a fixed position angle $\theta = 315°$. The dashed lines correspond to no wavefront error while the solid lines correspond to 16 nm rms of wavefront error. We can see that the theoretical performance, validated for a single companion signature in Fig. 7, hold true over a large range of contrasts and separations, and that the detection limit of T_{B} remains close to the bound provided by T_{NP}. The dashed and dotted lines correspond to a perfectly stable JWST leading to a perfect calibration of the systematic errors.

The detection limits further depend on θ, because the PSF of JWST NIRISS is not centrosymmetric (as visible in Fig. 1). Fluctuations of these limits are shown in Fig. 9 for three position angles. Figure 9 also indicates the S/N level at the corresponding positions in the image (computed here as the maximal pixel value of a noiseless image with only the companion, divided by the standard deviation of the considered noise), showing that the detection limits follow the overall noise level in the image. Performance wise, the detectable contrast ratios are of the order of 10^3 at 200 mas, with some variations between the two flux levels considered.

[5] The number of realisations is dictated by the target P_{FA} and P_{DET}. For the considered $P_{\mathrm{FA}} = 1\%$ and $P_{\mathrm{DET}} = 68\%$, 2000 realisations correctly sample the distributions of the test statistic of T_{B} under $\mathcal{H}_0$ and $\mathcal{H}_1$.

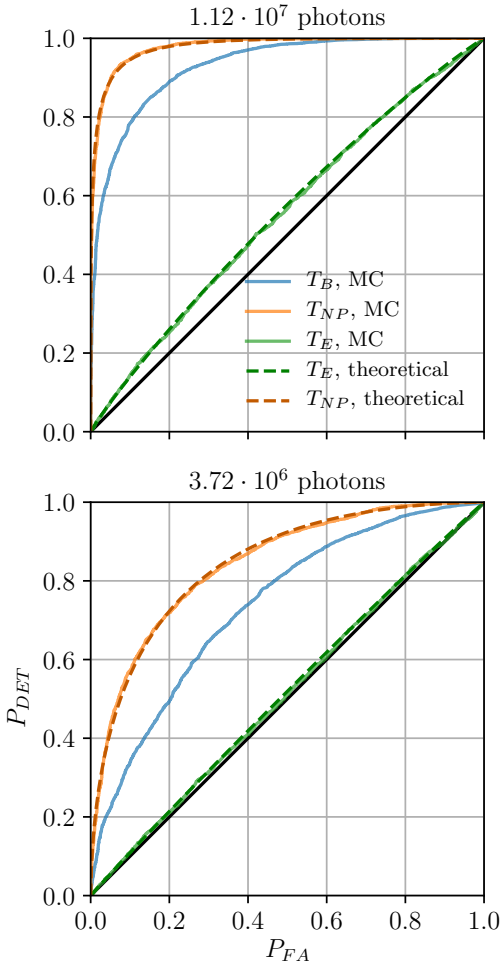

Fig. 7. ROC curves of T_{E} (green), T_{NP} (blue), and T_{B} (orange). Theoretical ROC curves for T_{NP} and T_{E} plotted using Eqs. (17) and (24), for a companion at $\rho = 200$ mas, $c = 1200$, and $\theta = 45°$ off the vertical. Dashed lines correspond to theoretical ROCs, while solid lines represent ROCs obtained by Monte-Carlo simulations. The closer a curve is to the black line on the diagonal, the less powerful the corresponding test. The higher flux regime is represented in the *top panel*, and the lower flux regime in the *bottom panel*. The performance of T_{NP} and T_{E} are accurately described by the theoretical expressions in Eqs. (17) and (24). The test T_{NP} presents the highest performance. T_{B} is the next-best-performing test and T_{E} has the lowest performance of the three. We see a clear improvement of the power of all tests as the flux (and thus the S/N) increases.

3.6. Mass limits for WISE 1405+5534

WISE 1405+5534 is a Y-type brown dwarf with a W2 magnitude of 14.1 that was used as a reference target to produce the contrast detection limits featured in Fig. 8. The raw observational detection limit curve of contrast as a function of angular separation can be converted into an astrophysical detection limit curve of companion mass as a function of orbital separation.

Whereas the 129 ± 19 mas parallax measured by Dupuy & Kraus (2013) directly allows for the conversion of the angular separation into a projected orbital distance, the contrast to mass conversion requires a model. We use the mass–luminosity relations given by the AMES-Cond model of Baraffe et al. (2003) for an age of 1 Gyr and a mass estimate of 30 M_J for the primary given by Cushing et al. (2011).

The detection limits obtained for WISE 1405+5534 are shown in Fig. 10. At $P_{\mathrm{FA}} = 1\%$, and $P_{\mathrm{DET}} = 68\%$, a 1 M_J can

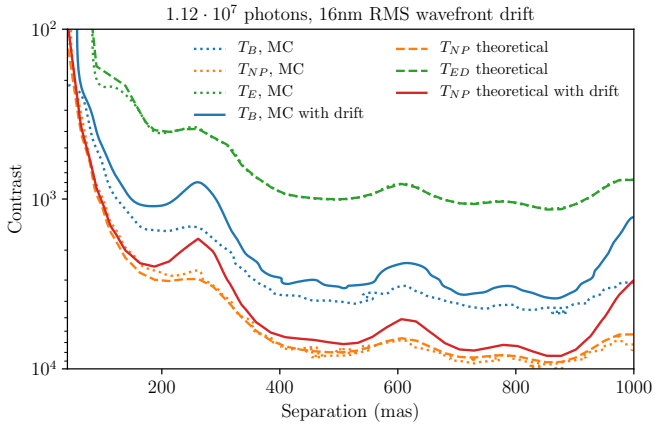

Fig. 8. Detection limits at a fixed position angle $\theta = 315°$: below the contours the $P_{\rm DET}$ falls below 68% for a fixed $P_{\rm FA}$ of 1%, represented as a function of the separation and contrast of the companions, for $T_{\rm E}$ (green), $T_{\rm B}$ (blue), and $T_{\rm NP}$ (orange). The dashed lines represent theoretical detection limits for $T_{\rm E}$ and $T_{\rm NP}$ (Eqs. (17) and (24)) and the dotted lines present the limits achieved in the MC simulations. $T_{\rm NP}$ (orange) provides ideal detection limits for a Kernel treatment of a JWST-NIRISS image and the practical test $T_{\rm B}$ (dotted blue) has contrast detection limits within a factor of 2.5 of the theoretical maximum. The solid lines represent the detection limits for $T_{\rm B}$ (blue) and $T_{\rm NP}$ (red) with a calibration residual corresponding to a 16 nm rms wavefront drift.

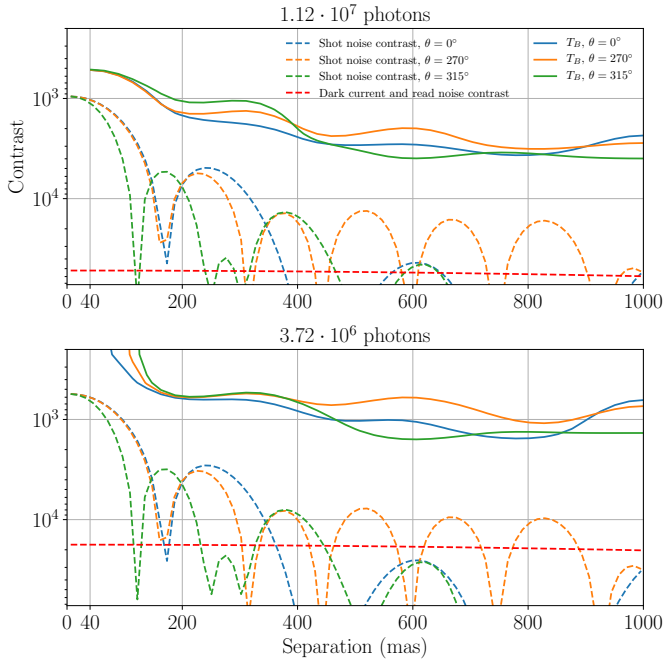

Fig. 9. Detection limits for test $T_{\rm B}$ (Eq. (31)), in the higher flux regime (*top panel*) and the lower flux regime (*bottom panel*). The solid lines correspond to contours of $P_{\rm DET} = 68\%$ at a fixed $P_{\rm FA} = 1\%$. Detection limits are represented at three different position angles for the companion: 0, 45, and 90° off the vertical, as orientated in the PSF shown in Fig. 1. The relative S/Ns (see text) are indicated by dashed lines. The shot (photon) noise is the main limiting noise in most cases.

be detected at separations greater that 1.5 AU. An orbit with this semi major axis would have a period of 40 years, thus a quarter of an orbit could be captured with repeated observations over the expected service life of JWST.

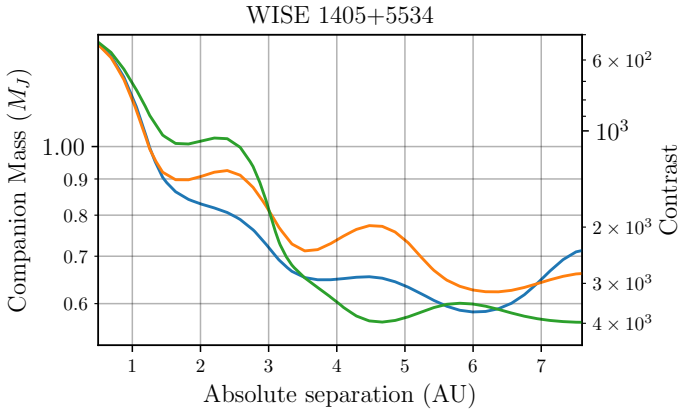

Fig. 10. Detection limits of a possible companion to WISE 1405+5534 at $P_{\rm FA} = 1\%$ and $P_{\rm DET} = 68\%$, as a function of contrast (right ordinate axis) or mass (left ordinate axis) and absolute separation in AU. A one-Jupiter-mass object is detectable down to 1.5 AU from the primary.

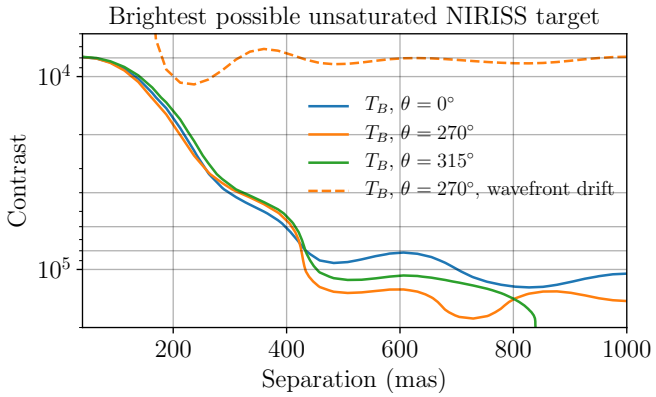

Fig. 11. Detection limits for the brightest target observable without saturation with JWST NIRISS. Solid lines show detection limits for $T_{\rm B}$ at $P_{\rm FA} = 1\%$ and $P_{\rm DET} = 68\%$ applied to the image with the greatest possible dynamic range, with 20 min total integration time. For the brightest images, the kernel method with the test $T_{\rm B}$ ideally allows detection of contrasts up to 10^5 beyond 500 mas. The dashed orange line represents detection limits in the presence of a 16 nm wavefront drift.

3.7. Bright limits

For the faint Y-dwarf targets considered thus far, it may have occurred to the reader that the contrast detection limits are dominated by the effect of the dark current and the readout noise and not by the photon noise of the central object. We wish here to complete the description of the properties of our approach with a bright target scenario that will feature a different behaviour, thus exhibiting the contribution of the photon noise.

The saturation limit for full-pupil JWST NIRISS using the F480M filter and a 64×64 pixels subarray size is 7.6 mag. We consider a shorter observation sequence, with a total of 20 min spent on the target of interest and 20 min on a calibrator of similar brightness. The detection limits for this observation using the operational test $T_{\rm B}$ are shown in Fig. 11, at $P_{\rm FA} = 1\%$, and $P_{\rm DET} = 68\%$.

Unlike the contrast detection limits obtained on the faint targets, the curves now clearly reveal two different regimes. Up to an angular separation of ≈500 mas, where the photon noise is expected to dominate, the contrast detection decreases as a function of the angular separation. Beyond this point, it reaches a plateau, as the detection is once again dominated by

the homogeneous properties of the dark current and the readout noise.

In this bright scenario, calibration errors induced by a drift comparable to what was described in Sect. 3.5 will have a stronger impact on the weak signal of a high-contrast companion. Sallum & Skemer (2019) feature contrast detection limits for NIRCam in a similar scenario that takes calibration errors into account. Under the hypothesis introduced in Sect. 3.2, the calibration error accounts here for 85% of the total noise variance of the kernels and therefore results in a degraded performance by a factor of approximately 10, as shown by the dashed curve in Fig. 11.

4. Conclusion

This paper provides a theoretical and numerical analysis of the performance of various detection tests based on the Kernel method. The approach provides an upper bound for the achievable detection limits, and an operational detection test whose performances are close to the upper bound. Furthermore, the false-alarm rate of these tests is not affected by fluctuating aberrations and can be tuned a priori.

The kernel-based detection approach presented in this paper is not specific to either NIRISS, the 480M band, the full-pupil imaging mode, or to JWST itself. The method only requires weak wavefront perturbations and appropriate sampling (i.e. a small-enough plate scale as compared to λ/D). In particular, the statistical treatment proposed in this study can also be used for NRM data.

For JWST-NIRISS in the F480M band, we have shown that medium-($\approx 10^2$) to high-($\approx 10^3$)contrast detections can realistically be achieved for separations down to half of λ/D on ultracool brown dwarf primary targets. In practice, this means that a 80 min observation sequence can allow for the detection of a $1\,M_J$ situated 1.5 AU away from a $30\,M_J$ Y-type brown dwarf at a distance of 8 pc. On brighter targets, kernel-phase analysis combined with the methods presented in this paper can reveal companions at contrasts $\approx 10^3$ down to $0.3\,\lambda/D$.

Detection results presented in this paper rely on up-to-date simulations of JWST-NIRISS frames that take into account all the noises expected to contribute to kernel-phase uncertainties. These results can be affected by several factors that are not yet accounted for, the most critical being probably calibration errors. Instrumental drifts in the range of a few tens of nanometres, as predicted by Perrin et al. (2018), are not expected to significantly degrade performances for Y dwarfs. Another limitation may come from the algorithmic efficiency in determining the MLE $\widehat{\boldsymbol{x}}$ in Eq. (29) for the test T_{B}. Overly coarse grid searches or algorithms too sensitive to local minima will lead to a loss in detection power and to an increased uncertainty for the estimated parameters.

The performance reported in this work can therefore be seen as ideal contrast performance achievable using kernel phases for JWST NIRISS images. The method can in principle be improved upon by exploiting the full information available in the image (present not only in the phase but also, to a lesser extent, in the amplitude of the complex visibility). Even working solely with the phase, the calibration problem can be mitigated by using a more accurate and less idealised representation of the instrument. A significant fraction of the calibration error comes from the use of a necessarily approximate discrete model to represent the continuous phenomenon of diffraction. The results reported in this work were achieved using a dense aperture model to mitigate this discretisation error; however, the representation is not yet optimal. One avenue to improve the overall fidelity for example seems to be to take into account a variable local transmission function to more accurately describe the aperture with the same grid density. The study of the general aperture modelling prescription will be the object of future work.

The `XARA` package is regularly updated in the context of the KERNEL project.

Acknowledgements. KERNEL has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement CoG – #683029). We wish to thank Nick Cvetojevic for proof-reading the paper and his contribution to group discussions. The manuscript was also significantly improved following the comments and recommendations from the anonymous referee.

References

Andrieu, C., de Freitas, N., Doucet, A., & Jordan, M. I. 2003, Mach. Learn., 50, 5
Artigau, Ă., Sivaramakrishnan, A., Greenbaum, A. Z., et al. 2014, SPIE Conf. Ser., 9143, 914340
Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Nature, 320, 595
Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701
Bernat, D., Bouchez, A. H., Ireland, M., et al. 2010, ApJ, 715, 724
Branch, M. A., Coleman, T. F., & Li, Y. 1999, SIAM J. Sci. Comput., 21, 442
Cushing, M. C., Kirkpatrick, J. D., Gelino, C. R., et al. 2011, ApJ, 743, 50
Dupuy, T. J., & Kraus, A. L. 2013, Science, 341, 1492
Fontanive, C., Biller, B., Bonavita, M., & Allers, K. 2018, MNRAS, 479, 2702
Ford, K. E., McKernan, B., Sivaramakrishnan, A., et al. 2014, AJ, 783, A73
Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Sci. Rev., 123, 485
Greenbaum, A. Z., Pueyo, L., Sivaramakrishnan, A., & Lacour, S. 2015, ApJ, 798, 68
Greenbaum, A., Sivaramakrishnan, A., Sahlmann, J., & Thatte, D. 2016, `https://github.com/agreenbaum/ami_sim`
Greenbaum, A. Z., Pueyo, L., Ruffio, J.-B., et al. 2018, AJ, 155, 226
Huélamo, N., Lacour, S., Tuthill, P., et al. 2011, A&A, 528, L7
Ireland, M. J. 2013, MNRAS, 433, 1718
Jennison, R. C. 1958, MNRAS, 118, 276
Kammerer, J., Ireland, M. J., Martinache, F., & Girard, J. H. 2019, MNRAS, 486, 639
Kirkpatrick, J. D., Cushing, M. C., Gelino, C. R., et al. 2011, ApJS, 197, 19
Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762
Kraus, A. L., Ireland, M. J., Martinache, F., & Hillenbrand, L. A. 2011, ApJ, 731, 8
Lacour, S., Tuthill, P., Amico, P., et al. 2011, A&A, 532, A72
Laugier, R., Martinache, F., Ceau, A., et al. 2019, A&A, 623, A164
Le Bouquin, J.-B., & Absil, O. 2012, A&A, 541, A89
Marsh, K. A., Wright, E. L., Kirkpatrick, J. D., et al. 2013, ApJ, 762, 119
Martinache, F. 2010, ApJ, 724, 464
Martinache, F., Guyon, O., Garrel, V., et al. 2009, AIP Conf. Proc., 1158, 393
Nakajima, T. 1989, AJ, 97, 1510
Neyman, J., & Pearson, E. S. 1933, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci., 231, 289
Perrin, M. D., Pueyo, L., Rajan, A., et al. 2018, in Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, eds. H. A. MacEwen, M. Lystrup, G. G. Fazio, et al. (SPIE), 1069809, 8
Pope, B., Martinache, F., & Tuthill, P. 2013, ApJ, 767, 110
Pravdo, S. H., Shaklan, S. B., Wiktorowicz, S. J., et al. 2006, ApJ, 649, 389
Sallum, S., & Skemer, A. 2019, J. Astron. Telesc. Instrum. Syst., 5, 1
Sallum, S., Follette, K. B., Eisner, J. A., et al. 2015, Nature, 527, 342
Scharf, L., & Friedlander, B. 1994, IEEE Trans. Sign. Process., 42, 2146
Schneider, A. C., Cushing, M. C., Kirkpatrick, J. D., et al. 2015, ApJ, 804, 92
Sivaramakrishnan, A., Lafrenière, D., Ford, K. E. S., et al. 2012, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, 8442, 84422S
Skilling, J. 2004, AIP Conf. Proc., 735, 395
STSCI 2018, NIRISS Imaging, Space Telescope Science Institute, James Webb Space Telescope User Documentation
Tuthill, P., Monnier, J., Tanner, A., et al. 2006, Science, 313, 935
van Cittert, P. 1934, Physica, 1, 201
Wright, E. L., Eisenhardt, P. R., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Zernike, F. 1938, Physica, 5, 785
Zwieback, S., Liu, X., Antonova, S., et al. 2016, IEEE Trans. Geosci. Remote Sens., 54, 5588

Bibliography

Airy, G. B. (1835). "On the Diffraction of an Object-glass with Circular Aperture". In: *Transactions of the Cambridge Philosophical Society*, 5, p. 283.

Baraffe, I et al. (2003). "Evolutionary models for cool brown dwarfs and extrasolar giant planets. The case of HD 209458". In: *Astronomy & Astrophysics* 402.2, pp. 701–712. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20030252. URL: http://arxiv.org/abs/astro-ph/0302293http://dx.doi.org/10.1051/0004-6361:20030252http://www.aanda.org/10.1051/0004-6361:20030252.

Burrows, Adam et al. (2001). "The theory of brown dwarfs and extrasolar giant planets". In: *Reviews of Modern Physics* 73.3, pp. 719–765. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.73.719. arXiv: 0103383v1 [astro-ph]. URL: https://link.aps.org/doi/10.1103/RevModPhys.73.719.

Cantalloube, F. et al. (2015). "Direct exoplanet detection and characterization using the ANDROMEDA method: Performance on VLT/NaCo data". In: *Astronomy & Astrophysics* 582, A89. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201425571. URL: http://www.aanda.org/10.1051/0004-6361/201425571.

Ceau, A. et al. (2019). "Kernel-phase detection limits". In: *Astronomy & Astrophysics* 630, A120. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201935247. URL: https://www.aanda.org/10.1051/0004-6361/201935247.

Chauvin, G. et al. (2004). "A giant planet candidate near a young brown dwarf". In: *Astronomy & Astrophysics* 425.2, pp. L29–L32. ISSN: 0004-6361. DOI: 10.1051/0004-6361:200400056. URL: http://arxiv.org/abs/astro-ph/0409323%0Ahttp://dx.doi.org/10.1051/0004-6361:200400056http://www.aanda.org/10.1051/0004-6361:200400056.

Cittert, P.H van (1934). "Die Wahrscheinliche Schwingungsverteilung in Einer von Einer Lichtquelle Direkt Oder Mittels Einer Linse Beleuchteten Ebene". In: *Physica* 1.1-6, pp. 201–210. ISSN: 00318914. DOI: 10.1016/S0031-8914(34)90026-4. URL: http://linkinghub.elsevier.com/retrieve/pii/S0031891434900264.

Cushing, Michael (2019). *Y Dwarf Compendium*. URL: https://sites.google.com/view/ydwarfcompendium/.

Cushing, Michael C et al. (2011). "The Discovery of Y Dwarfs Using Data from the Wide-field Infrared Survey Explorer (WISE)". In: *The Astrophysical Journal* 743.1, p. 50. ISSN: 0004-637X. DOI: 10.1088/0004-637X/743/1/50. URL: http://stacks.iop.org/0004-637X/743/i=1/a=50?key=crossref.58412f84087d0af7c5ff252457e8cf30http://arxiv.org/abs/1108.4678http://dx.doi.org/10.1088/0004-637X/743/1/50.

Cutri, R. M. et al. (2015). *Explanatory Supplement to the NEOWISE Data Release Products*. URL: http://wise2.ipac.caltech.edu/docs/release/neowise/expsup.

Dohlen, Kjetil et al. (2006). "SPHERE: A planet finder instrument for the VLT". In: ed. by Ian S. McLean and Masanori Iye, 62690Q. DOI: 10.1117/12.671537. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.671537.

Doyon, René et al. (2012). "The JWST Fine Guidance Sensor (FGS) and Near-Infrared Imager and Slitless Spectrograph (NIRISS)". In: *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. Proceedings of the SPIE*. Ed. by Mark C. Clampin et al. Vol. 8442, 84422R. DOI: 10.1117/12.926578. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.926578.

Dupuy, Trent J. and Adam L. Kraus (2013). "Distances, luminosities, and temperatures of the coldest known substellar objects". In: *Science* 341.6153, pp. 1492–1495. ISSN: 10959203. DOI: 10.1126/science.1241917.

Fizeau, Hyppolyte (1868). "Rapport sur le concours du Prix Bordin de l'année 1867". In: *Comptes Rendus des Séances de l'Académie des Sciences* 66, pp. 932–934.

Fontanive, Clémence et al. (2018). "Constraining the multiplicity statistics of the coolest brown dwarfs: binary fraction continues to decrease with spectral type". In: *Monthly Notices of the Royal Astronomical Society* 479.2, pp. 2702–2727. ISSN: 0035-8711. DOI: 10.1093/mnras/sty1682. URL: http://arxiv.org/abs/1806.08737http://dx.doi.org/10.1093/mnras/sty1682https://academic.oup.com/mnras/article/479/2/2702/5045265.

Fresnel, Augustin (1823). "Mémoire sur la diffraction de la lumière". In: *Mémoires de l'Académie des sciences de l'Institut de France* 5, 339–475.

Fried, D. L. (1966). "Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures". In: *Journal of the Optical Society of America* 56.10, p. 1372. ISSN: 0030-3941. DOI: 10.1364/JOSA.56.001372. URL: https://www.osapublishing.org/abstract.cfm?URI=josa-56-10-1372.

Gauchet, L. et al. (2016). "Sparse aperture masking at the VLT". In: *Astronomy & Astrophysics* 595, A31. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201526404. URL: http://www.aanda.org/10.1051/0004-6361/201526404.

Ghez, Andrea M. et al. (1990). "Infrared speckle imaging at Palomar." In: ed. by James B. Breckinridge, pp. 249–258. DOI: 10.1117/12.19298. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=938615.

Goodman (1986). "Statistical optics". In: 22.5, pp. 347–355. ISSN: 0018-9197. DOI: 10.1109/JQE.1986.1073023. URL: http://ieeexplore.ieee.org/document/1073023/.

Greenbaum, Alexandra Z., Anand Sivaramakrishnan, and Laurent Pueyo (2013). "Flat field errors and intra-pixel sensitivities for non-redundant aperture masking interferometry on JWST NIRISS". In: ed. by Stuart Shaklan, p. 88641L. DOI: 10.1117/12.2023719. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2023719.

Greenbaum, Alexandra Z. et al. (2014). "Analyzing the first JWST-NIRISS NRM test data". In: ed. by Jacobus M. Oschmann et al., p. 91434M. DOI: 10.1117/12.2056346. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2056346.

Hanbury Brown, R. and R. Q. Twiss (1956). "A Test of a New Type of Stellar Interferometer on Sirius". In: *Nature* 178.4541, pp. 1046–1048. ISSN: 0028-0836. DOI: 10.1038/1781046a0. URL: http://www.nature.com/articles/1781046a0.

Herschel, John (1828). "Treatises on physical astronomy, light and sound contributed to the Encyclopaedia metropolitana". In: *Encyclopaedia metropolitana*. London, R. Griffin. Chap. Treatises, p. 491.

Howard, Joseph M. and Lee D. Feinberg (2009). "Optical modeling activities for NASA's James Webb Space Telescope (JWST): VI. secondary mirror figure compensation using primary mirror segment motions". In: ed. by Howard A. MacEwen and James B. Breckinridge, p. 74360C. DOI: 10.1117/12.825420. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.825420.

Ireland, M. J. (2013). "Phase errors in diffraction-limited imaging: Contrast limits for sparse aperture masking". In: *Monthly Notices of the Royal Astronomical Society* 433.2, pp. 1718–1728. ISSN: 00358711. DOI: 10.1093/mnras/stt859.

Jennison, R. C. (1958). "A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent". In: *Monthly Notices of the Royal Astronomical Society* 118, p. 276. DOI: 10.1093/mnras/118.3.276.

Johnston, John D. et al. (2004). "Integrated modeling activities for the James Webb Space Telescope: structural-thermal-optical analysis". In: *Optical, Infrared, and Millimeter Space Telescopes* 5487, p. 600. DOI: 10.1117/12.551704.

Kammerer, Jens et al. (2019). "Kernel phase imaging with VLT/NACO: high-contrast detection of new candidate low-mass stellar companions at the diffraction limit". In: *Monthly Notices of the Royal Astronomical Society* 486.1, pp. 639–654. ISSN: 0035-8711. DOI: 10.1093/mnras/stz882. URL: https://academic.oup.com/mnras/article/486/1/639/5421634.

Käufl, Hans-Ulrich et al. (2018). "NEAR: New Earths in the Alpha Cen Region (bringing VISIR as a "visiting instrument" to ESO-VLT-UT4)". In: July, p. 12. ISSN: 1996756X. DOI: 10.1117/12.2313395.

Kooten, M. A. M. van, N. Doelman, and M. Kenworthy (2020). "Robustness of prediction for extreme adaptive optics systems under various observing conditions". In: *Astronomy & Astrophysics* 636, A81. ISSN: 0004-6361. DOI:

10.1051/0004-6361/201937076. URL: https://www.aanda.org/10.1051/0004-6361/201937076.

Labeyrie, A (1970). "Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier Analysing Speckle Patterns in Star Images". In: *Astronomy and Astrophysics* 6.1, pp. 85–87. ISSN: 0004-6361. URL: http://adsabs.harvard.edu/abs/1970A%26A.....6...85L%5Cnpapers3://publication/uuid/02B091DE-642D-4FA3-8BF8-3D83E11B27F2.

Labeyrie, A. (1975). "Interference fringes obtained on VEGA with two optical telescopes". In: *The Astrophysical Journal* 196, p. L71. ISSN: 0004-637X. DOI: 10.1086/181747. URL: http://adsabs.harvard.edu/doi/10.1086/181747.

Labeyrie, Antoine et al. (2001). "Hypertelescopes: Prospects on Earth and in Space". In: *SF2A-2001: Semaine de l'Astrophysique Francaise, meeting held in Lyon, France, May 28-June 1st, 2001, Eds.: F. Combes, D. Barret, F. Thévenin*. Lyon.

Lagarde, S. et al. (2006). "MATISSE: a four beams combiner in the mid-infrared for the VLTI". In: ed. by John D. Monnier, Markus Schöller, and William C. Danchi, p. 62683M. DOI: 10.1117/12.671690. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.671690.

Laugier, Romain et al. (2019a). "Angular Differential Kernel-phase". In: *Astronomy and Astrophysics (Submitted 2019)*.

Laugier, Romain et al. (2019b). "Recovering saturated images for high dynamic Kernel-Phase analysis Application to the determination of dynamical masses for the system Gl 494AB". In: *Astronomy & Astrophysics* 494. URL: http://arxiv.org/abs/1901.02824.

Lawrence, Jon S. et al. (2004). "Exceptional astronomical seeing conditions above Dome C in Antarctica". In: *Nature* 431.7006, pp. 278–281. ISSN: 0028-0836. DOI: 10.1038/nature02929. URL: http://www.nature.com/articles/nature02929.

Leggett, S. K. et al. (2017). "The Y-Type Brown Dwarfs: Estimates of Mass and Age from New Astrometry, Homogenized Photometry and Near-Infrared Spectroscopy". In: *The Astrophysical Journal* 842.2, p. 118. ISSN: 1538-4357. DOI: 10.3847/1538-4357/aa6fb5. URL: http://arxiv.org/abs/1704.03573%0Ahttp://dx.doi.org/10.3847/1538-4357/aa6fb5http://arxiv.org/abs/1704.03573http://dx.doi.org/10.3847/1538-4357/aa6fb5.

Lehmann, Erich L. and Joseph P. Romano (1959). *Testing Statistical Hypotheses*. 2008th ed., pp. 60–65. ISBN: 978-0-387-27605-2. DOI: 10.1177/019262338901700403.

Lohmann, Adolf W., Gerd Weigelt, and Bernhard Wirnitzer (1983). "Speckle masking in astronomy: triple correlation theory and applications". In: *Applied Optics* 22.24, p. 4028. ISSN: 0003-6935. DOI: 10.1364/AO.22.004028. URL: https://www.osapublishing.org/abstract.cfm?URI=ao-22-24-4028.

Lozi, Julien et al. (2018). "SCExAO, an instrument with a dual purpose: perform cutting-edge science and develop new technologies". In: *Adaptive*

Optics Systems VI. Ed. by Dirk Schmidt, Laura Schreiber, and Laird M. Close. SPIE, p. 270. ISBN: 9781510619593. DOI: 10.1117/12.2314282. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10703/2314282/SCExAO-an-instrument-with-a-dual-purpose--perform-cutting/10.1117/12.2314282.full.

Macintosh, Bruce et al. (2006). "The Gemini Planet Imager". In: p. 62720L. DOI: 10.1117/12.672430. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.672430.

Marcum, J. I. (1947). *A Statistical Theory of Target Detection by Pulsed Radar*. Santa Monica, CA: RAND Corporation.

Marois, Christian et al. (2006). "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique". In: *The Astrophysical Journal* 641.1, pp. 556–564. ISSN: 0004-637X. DOI: 10.1086/500401. arXiv: 0512335v1 [astro-ph]. URL: http://stacks.iop.org/0004-637X/641/i=1/a=556.

Martinache, F. (2018). "Repousser les limites de la diffraction pour l'astronomie à haute résolution angulaire". PhD thesis. Université de Nice Sophia Antipolis.

Martinache, Frantz (2010). "Kernel Phase in Fizeau Interferometry". In: *The Astrophysical Journal* 724.1, pp. 464–469. ISSN: 0004-637X. DOI: 10.1088/0004-637X/724/1/464. URL: http://arxiv.org/abs/1009.3933%0Ahttp://dx.doi.org/10.1088/0004-637X/724/1/464http://arxiv.org/abs/1009.3933http://dx.doi.org/10.1088/0004-637X/724/1/464http://stacks.iop.org/0004-637X/724/i=1/a=464?key=crossref.aecd5d349fbbfdc162ccab1b74b2b5d5https:/.

Martinache, Frantz et al. (2020). "Kernel-phase analysis: aperture modeling prescriptions that minimize calibration errors". In: *Astronomy & Astrophysics*.

McAlister, Harold A. et al. (1995). "CHARA optical/IR interferometric array project". In: ed. by William H. Carter, pp. 180–188. DOI: 10.1117/12.219572. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1004238.

Mccord, Thomas B and Jeffrey P Bosel (1975). "Potential usefulness of CCD imagers in astronomy". In: pp. 65–69. ISBN: 1975002075. URL: https://ntrs.nasa.gov/search.jsp?R=19750020756.

McPherson, Alistair et al. (2012). "E-ELT update of project and effect of change to 39m design". In: ed. by Larry M. Stepp, Roberto Gilmozzi, and Helen J. Hall, 84441F. DOI: 10.1117/12.927787. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.927787.

Michelson, A. A. and F. G. Pease (1921). "Measurement of the diameter of alpha Orionis with the interferometer." In: *The Astrophysical Journal* 53.203, p. 249. ISSN: 0004-637X. DOI: 10.1086/142603. URL: http://adsabs.harvard.edu/doi/10.1086/142603.

MIRI Filters and Dispersers. URL: https://jwst-docs.stsci.edu/mid-infrared-instrument/miri-instrumentation/miri-filters-and-dispersers.

Nakajima, T. et al. (1989). "Diffraction-limited imaging. II - Optical aperture-synthesis imaging of two binary stars". In: *The Astronomical Journal* 97.5, p. 1510. ISSN: 00046256. DOI: 10.1086/115090. URL: http://adsabs.harvard.edu/abs/1989AJ.....97.1510Nhttp://adsabs.harvard.edu/cgi-bin/bib_query?1989AJ.....97.1510N.

Neyman, J. and E. S. Pearson (1933). "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 231.694-706, pp. 289–337. ISSN: 1364-503X. DOI: 10.1098/rsta.1933.0009. URL: http://rsta.royalsocietypublishing.org/cgi/doi/10.1098/rsta.1933.0009.

NIRCam Filters (2019). URL: https://jwst-docs.stsci.edu/near-infrared-camera/nircam-instrumentation/nircam-filters.

NIRISS Filters (2019). URL: https://jwst-docs.stsci.edu/near-infrared-imager-and-slitless-spectrograph/niriss-instrumentation/niriss-filters.

Oliveira, Catarina Alves de et al. (2018). "Preparing the NIRSpec/JWST science data calibration: from ground testing to sky". In: URL: http://arxiv.org/abs/1805.06922.

Perrin, Marshall D. et al. (2018). "Updated optical modeling of JWST coronagraph performance contrast, stability, and strategies". In: *Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave*. Ed. by Howard A. MacEwen et al. Vol. 1069809. August 2018. SPIE, p. 8. ISBN: 9781510619494. DOI: 10.1117/12.2313552. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10698/2313552/Updated-optical-modeling-of-JWST-coronagraph-performance-contrast-stability-and/10.1117/12.2313552.full.

Pope, Benjamin, Frantz Martinache, and Peter Tuthill (2013). "Dancing in the Dark: New Brown Dwarf Binaries From Kernel Phase Interferometry". In: *The Astrophysical Journal* 767.2, p. 110. ISSN: 0004-637X. DOI: 10.1088/0004-637X/767/2/110. URL: http://stacks.iop.org/0004-637X/767/i=2/a=110?key=crossref.55fdaeb3902008d8ed564752c8da0466.

Pope, Benjamin et al. (2016). "The Palomar kernel-phase experiment: Testing kernel phase interferometry for ground-based astronomical observations". In: *Monthly Notices of the Royal Astronomical Society* 455.2, pp. 1647–1653. ISSN: 13652966. DOI: 10.1093/mnras/stv2442.

Pope, Benjamin J. S. (2016). "Kernel phase and kernel amplitude in Fizeau imaging". In: *Monthly Notices of the Royal Astronomical Society* 463.4, pp. 3573–3581. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2215. URL: https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stw2215.

Rayleigh (1879). "XXXI. Investigations in optics, with special reference to the spectroscope". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 8.49, pp. 261–274. ISSN: 1941-5982. DOI: 10.1080/14786447908639684. URL: https://www.tandfonline.com/doi/full/10.1080/14786447908639684.

Sallum, Steph and Andy Skemer (2019). "Comparing nonredundant masking and filled-aperture kernel phase for exoplanet detection and characterization". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 5.01, p. 1. ISSN: 2329-4124. DOI: 10.1117/1.JATIS.5.1.018001. URL: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes-Instruments-and-Systems/volume-5/issue-01/018001/Comparing-nonredundant-masking-and-filled-aperture-kernel-phase-for-exoplanet/10.1117/1.JATIS.5.1.018001.full.

Scharf, L.L. and Benjamin Friedlander (1994). "Matched subspace detectors". In: *IEEE Transactions on Signal Processing* 42.8, pp. 2146–2157. ISSN: 1053587X. DOI: 10.1109/78.301849. URL: http://ieeexplore.ieee.org/document/301849/.

Sivaramakrishnan, Anand et al. (2012). "Non-redundant Aperture Masking Interferometry (AMI) and segment phasing with JWST-NIRISS". In: ed. by Mark C. Clampin et al., 84422S. DOI: 10.1117/12.925565. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.925565.

Solomon, Daniel L. (1975). "A Note on the Non-equivalence of the Neyman-Pearson and Generalized Likelihood Ratio Tests for Testing a Simple Null versus a Simple Alternative Hypothesis". In: *The American Statistician* 29.2, pp. 101–102. ISSN: 0003-1305. DOI: 10.1080/00031305.1975.10477383. URL: http://www.tandfonline.com/doi/abs/10.1080/00031305.1975.10477383.

Station, Bank Experimental (1957). "Interferometry of the intensity fluctuations in light - I. Basic theory: the correlation between photons in coherent beams of radiation". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 242.1230, pp. 300–324. ISSN: 2053-9169. DOI: 10.1098/rspa.1957.0177. URL: http://www.royalsocietypublishing.org/doi/10.1098/rspa.1957.0177.

Stephan, Edouard (1874). "Sur l'extrême petitesse du diamètre apparent des étoiles fixes." In: *Comptes Rendus des Séances de l'Académie des Sciences. Paris.* Paris, pp. 1008–1012. URL: https://gallica.bnf.fr/ark:/12148/bpt6k3035z/f870.image.

STSCI (2018). *NIRISS Imaging*. URL: https://jwst-docs.stsci.edu/display/JTI/NIRISS+Imaging.

Subasavage, John P. et al. (2009). "The Solar Neighborhood. XXI. Parallax Results from the CTIOPI 0.9 m Program: 20 New Members of the 25 Parsec White Dwarf Sample". In: *The Astronomical Journal* 137.6, pp. 4547–4560. ISSN: 0004-6256. DOI: 10.1088/0004-6256/137/6/4547. URL: http://stacks.iop.org/1538-3881/137/i=6/a=4547?key=crossref.832148fbc93e4cf587fea455bfa46773https://iopscience.iop.org/article/10.1088/0004-6256/137/6/4547.

Tuthill, Peter G. (2012). "The unlikely rise of masking interferometry: leading the way with 19th century technology". In: *Optical and Infrared Interferometry III*. Ed. by Françoise Delplancke, Jayadev K. Rajagopal, and Fabien Malbet. Vol. 8445, p. 844502. ISBN: 9780819491466. DOI: 10.1117/12.925794. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.925794.

Usuda, Tomonori et al. (2014). "Preliminary design study of the TMT Telescope structure system: overview". In: ed. by Larry M. Stepp, Roberto Gilmozzi, and Helen J. Hall, 91452F. DOI: 10.1117/12.2055767. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2055767.

Young, Thomas (1823). *An introduction to medical literature*. Ed. by W. Phillips. Lodon.

Zernike, F. (1938). "The concept of degree of coherence and its application to optical problems". In: *Physica* 5.8, pp. 785–795. ISSN: 00318914. DOI: 10.1016/S0031-8914(38)80203-2. URL: http://linkinghub.elsevier.com/retrieve/pii/S0031891438802032.

Zhao, M. et al. (2008). "First Resolved Images of the Eclipsing and Interacting Binary β Lyrae". In: *The Astrophysical Journal* 684.2, pp. L95–L98. ISSN: 0004-637X. DOI: 10.1086/592146. URL: https://iopscience.iop.org/article/10.1086/592146.