



**HAL**  
open science

# Combining high-contrast imaging with high-resolution spectroscopy: Actual on-sky MIRI/MRS results compared to expectations

S. Martos, A. Bidot, A. Carlotti, D. Mouillet

## ► To cite this version:

S. Martos, A. Bidot, A. Carlotti, D. Mouillet. Combining high-contrast imaging with high-resolution spectroscopy: Actual on-sky MIRI/MRS results compared to expectations. *Astronomy & Astrophysics - A&A*, 2025, 697, pp.A27. <10.1051/0004-6361/202453382>. <hal-05057348>

**HAL Id: hal-05057348**

**<https://hal.science/hal-05057348v1>**

Submitted on 6 May 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Combining high-contrast imaging with high-resolution spectroscopy: Actual on-sky MIRI/MRS results compared to expectations

S. Martos<sup>1,\*</sup>, A. Bidot<sup>2</sup>, A. Carlotti<sup>1</sup>, and D. Mouillet<sup>1</sup>

<sup>1</sup> Institute of Planetology and Astrophysics of Grenoble (IPAG), University Grenoble Alpes, CNRS, Grenoble, France

<sup>2</sup> Space Telescope Science Institute (STScI), 3700 San Martin Drive, Baltimore, MD 21218, USA

Received 10 December 2024 / Accepted 2 April 2025

## ABSTRACT

**Context.** Combining high-contrast imaging with high-resolution spectroscopy represents a powerful approach to detecting and characterizing exoplanets around nearby stars, despite the challenges posed by their faintness. Instruments like VLT/SPHERE represent the state of the art in high-contrast imaging; however, their spectral resolution ( $R \approx 50$ ) limits them to basic characterization of close companions. These instruments can observe planets with masses as low as 5–10  $M_{\text{Jup}}$  at distances of around 10 AU from their stars. Detection limits are primarily constrained by speckle noise, which dominates over photon and detector noise at short separations around bright stars, even when advanced differential imaging techniques are used. Similarly, image stability also limits space-based high-contrast imaging capability. This speckle noise can, however, be largely mitigated by molecular mapping, a more recent method that leverages information from high-resolution spectroscopic data.

**Aims.** Our objective is to understand and predict the effective detection limits associated with spectro-imaging data after processing with molecular mapping. This involves analyzing the propagation of fundamental noise sources, such as photon and detector noise, and comparing these predictions to real instrument data to assess performance losses due to instrument-based factors. Our goal is to identify and propose potential mitigation strategies for these additional sources of noise. Another key aim is to compare the predictions made by our analytical approach with actual observational data to validate and refine the model's accuracy where necessary.

**Methods.** We analyzed JWST/MIRI/MRS data using the recently developed semi-analytical and numerical tool, FastCurves, and compared the results with outputs from the end-to-end MIRI simulator. This simulator allows one to examine nonideal instrumental effects in detail. Additionally, we applied principal component analysis (PCA), a statistical method that identifies correlated patterns in the data, to help isolate systematic effects, both with and without molecular mapping.

**Results.** Our analysis involves investigating the systematic effects introduced by the instrument, identifying their origins, and evaluating their impact on both noise and signal. We show that valuable insights are gained regarding the effects of straylight, fringes, and aliasing artifacts, each linked to different residual systematic noise terms in the data. The results are further supported by principal component analysis, which also demonstrates its effectiveness in mitigating these effects. Additionally, we explore the similarities and discrepancies between observed and modeled companion spectra from an astronomical perspective.

**Conclusions.** We modified FastCurves to account for systematic effects and improve its modeling of MIRI/MRS noise, with its signal-to-noise predictions validated against empirical data. In high-stellar-flux regimes, systematic noise imposes an ultimate contrast limit when using molecular mapping alone. Our methodology, demonstrated with MIRI/MRS data, could greatly benefit other instruments, aiding in the planning of observational programs. For future instruments like ELT/ANDES and ELT/PCS, it could also inform and guide their development.

**Key words.** techniques: high angular resolution – techniques: image processing – techniques: imaging spectroscopy

## 1. Introduction

The detection of exoplanets has revolutionized our understanding of the Universe and raised many fascinating questions. Their exhaustive characterization has three important underlying objectives: (1) to elucidate the complexities of their formation and early evolution by measuring the relative abundances of various molecules, which serve as a proxy for determining the physical processes at play (Oberg & Bergin 2016; Mollière et al. 2022); (2) to explore the rich diversity of planetary systems and configurations (Lissauer et al. 2011), allowing for comparisons within and between planetary systems, particularly in relation to our own Solar System; and (3) to determine the habitability of

Earth-like planets, and search for potential indicators of extraterrestrial biological activity (Turbet 2018; Wang et al. 2018).

The spectral characterization of the atmosphere of these planets is essential to achieve these objectives. Acquiring data with a sufficiently high signal-to-noise ratio (S/N) is challenging due to the observational properties of these planetary systems, the specifications of our telescopes and instruments, and, for ground-based observations, the presence of Earth's atmosphere, which spectrally filters incoming light and introduces short-lived wavefront aberrations.

The minimum planet-to-star flux ratio ranges from  $10^{-4}$  to  $10^{-6}$  for young giant planets and decreases to  $10^{-8}$  to  $10^{-10}$  (or even smaller) for older, smaller planets. In the first case, the photons received from the young planets are emitted during their cooling process, while in the second case, they mostly result from the star's light reflecting off the planet's surface. These

\* Corresponding author; [steven.martos@univ-grenoble-alpes.fr](mailto:steven.martos@univ-grenoble-alpes.fr)

low flux ratios would not be an issue if the angular separation between the planet and its host star were sufficiently large, considering our telescopes' size, and the wavelengths at which observations are made. Unfortunately, even for nearby stars, the angular separation is typically a fraction of an arcsecond, translating into a few units of wavelength over diameter.

While some observations do not require the star and planet to be angularly resolved, there are several advantages to resolving them. Unresolved observations at a high spectral resolution can leverage the relative radial velocity of the star and the planet to spectrally resolve the planet. However, the small planet-to-star flux ratio induces photon noise that strongly limits this type of observation. In the more favorable case of transiting planets, the change in the spectroscopic signal over time (before, during, and after the transit) can help in isolating the planetary contribution. However, transits occur over short durations, which limits the performance of this technique and underscores the need for instrument stability when using data obtained over multiple orbits. On the other hand, high-angular-resolution observations significantly reduce photon noise, albeit at the cost of increased instrument complexity: an extreme adaptive optics system is required to achieve diffraction-limited observations with ground-based telescopes, and a coronagraph, or other similar techniques such as dark-hole generation, to further lower photon noise.

High-contrast imaging, which spatially isolates most of the light of a host star from its planetary companions, has advanced significantly over the past two decades. Instruments such as the Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE; [Beuzit et al. 2019](#)) at the Very Large Telescope (VLT), the Gemini Planet Imager (GPI; [Macintosh et al. 2014](#)) at the Gemini South Telescope, and the Subaru Coronagraphic Extreme Adaptive Optics system (SCEAO; [Jovanovic et al. 2015](#)) at the Subaru Telescope have enabled the direct imaging of young (<100 Myr) planets with unprecedented spatial resolution. Despite these achievements, detecting exoplanets through high-contrast imaging remains challenging due to speckles (stellar light aberrations) that can mask faint planetary signals, particularly at close angular separations. Although various differential imaging techniques have been developed to extract companion signals from speckle noise, the latter still establishes the detection and characterization limits when searching for companions around bright stars at short separations.

To address this challenge, a powerful detection and characterization technique, called molecular mapping, has been developed. This pioneering method, first applied by [Hoeijmakers et al. \(2018\)](#), uses high-resolution spectroscopy to identify the distinct molecular signatures of a companion with a cooler atmosphere than its host star. As a result, the faint signals emitted by planets can be distinguished from their much brighter stars through cross-correlation techniques ([Konopacky et al. 2013](#)), even in the presence of speckles. This approach has already demonstrated its effectiveness with several instruments, including the Keck Planet Imager and Characterizer (KPIC; [Delorme et al. 2021](#)) at the Keck Observatory, the High-Resolution Imaging and Spectroscopy of Exoplanets (HiRISE; [Vigan et al. 2024](#)) at the VLT, and the Multi-Unit Spectroscopic Explorer (MUSE; [Jorquera et al. 2024](#)) also at the VLT.

The founding paper for the combination of high-contrast imaging and high spectral resolution by [Snellen et al. \(2015\)](#) introduced the idea that if direct imaging could achieve a contrast of  $10^{-3}$  and spectroscopy a contrast of  $10^{-4}$ , then their combination should theoretically offer a contrast of  $10^{-7}$ . The propagation of noise and the useful signal resulting from this

combination, considering post-processing techniques such as molecular mapping, has already been investigated by [Landman et al. \(2023\)](#) and [Bidot et al. \(2024\)](#). It was demonstrated that speckles are effectively filtered by this method, with detection becoming potentially limited by fundamental noise sources. However, this analysis has not yet been extended to real data to confirm whether fundamental noise indeed sets the limit. In this study, we test this assumption and investigate whether additional systematic effects further constrain detection. To achieve this, we employ FastCurves ([Bidot et al. 2024](#)), an analytical and numerical tool designed to predict integral field spectrograph (IFS) performance, which is used to estimate both fundamental and systematic noise contributions.

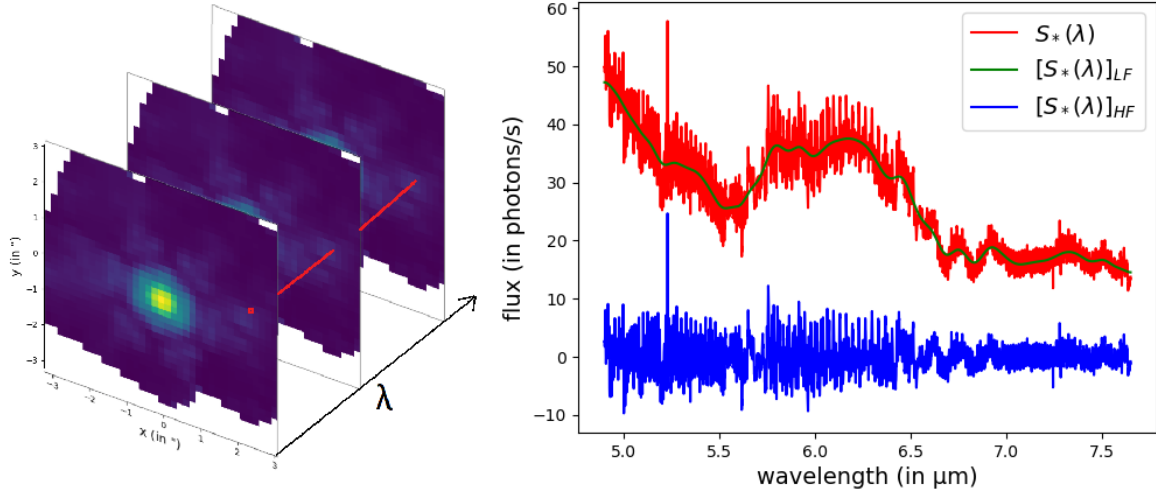
Currently, only a few IFSs provide diffraction-limited imaging on large telescopes with a high spectral resolution. In this regard, the medium-resolution spectroscopy mode of the Mid-IR Instrument (MIRI/MRS) on the *James Webb* Space Telescope (JWST) is particularly interesting, as it is not affected by adaptive optics residuals, making it simpler to study than ground-based instruments such as the Enhanced Resolution Imager and Spectrograph (ERIS; [Kravchenko et al. 2022](#)) on the VLT or the OH-Suppressing Infra-Red Integral-field Spectrograph (OSIRIS; [Petit dit de la Roche et al. 2018](#)) at the Keck Observatory. The proposed use of MIRI/MRS data, along with the MIRISim end-to-end simulator ([Klaassen et al. 2020](#)), offers a promising approach to understanding this data.

Firstly, we present the principles and formalism of molecular mapping, incorporating a new systematic term, along with the FastCurves tool, in Sect. 2. Then, we introduce MIRI/MRS, detailing its specifications and the method used for estimating systematic errors through end-to-end simulations (Sect. 3). Next, we detail an analysis of on-sky data, highlight the presence of systematic errors, and confirm the FastCurves noise model (Sect. 4). Following this, we identify the origins of these systematic errors, and we demonstrate how they impact the detection by altering noise statistics and by diminishing the planetary signal (Sect. 5). We also show how FastCurves can accurately estimate the planetary signal under these conditions and predict the potential mismatch between the observed planet spectrum and the considered model (Sect. 6). Finally, we show that principal component analysis (PCA) can be applied as a last resort to push the detection limits imposed by systematics, albeit with some reduction in signal (Sect. 7).

## 2. Methods and tools: Formalism and FastCurves

### 2.1. Molecular mapping principle

Speckles are intensity aberrations in starlight caused by wavefront errors that can evolve over time. When analyzing only photometric data, speckles can easily be mistaken for planetary signals and establish the detection limits for high-contrast imaging at close separations. Originating from diffraction effects ([Perrin et al. 2003](#)), speckles have a minimum size set by the wavelength-to-diameter ratio and a pattern that scales with wavelength at first order. As a result, the wavelength dependency of random speckle occurrences folds into the spectral dimension at a specific position, resulting in smooth (low-frequency) modulations within the spectra (see Fig. 1). High-pass filtering can be applied to remove these uncontrollable low-frequency modulations, which act as noise. At sufficiently high resolution ( $R > 100$ ), the spectral diversity between the star and the planet can be leveraged to distinguish and isolate the planetary flux.



**Fig. 1.** Left: MIRI/MRS data cube representation of a single star. Right: spectrum of the star at a given point  $(x, y)$  in red, with its modulated continuum in green and with high-pass filtering in blue.

The formalism that we present in this paper is based on Bidot et al. (2024). This framework quantifies (i) the high-pass filtered planetary signal relevant for detection (after accounting for its distinction from the stellar contribution, including potential self-subtraction when the stellar and planetary spectra share similarities) and (ii) the noise level due to fundamental sources. Observational data from spectro-imagers is represented as a 3D flux (data cube), which is a function of wavelength  $\lambda$  and spaxel (spatial pixel) position  $(x, y)$ . We let  $S$  denote a data cube (after reduction and calibration) corresponding to the measured flux (in electrons) integrated per pixel. It can be expressed as

$$\underbrace{S(\lambda, x, y)}_{\text{data}} = \underbrace{M(\lambda, x, y)\gamma(\lambda)S_*(\lambda)}_{\text{stellar halo}} + \underbrace{M_p(\lambda, x, y)\gamma(\lambda)S_p(\lambda)}_{\text{planet}} + \underbrace{n(\lambda, x, y)}_{\text{noise}}, \quad (1)$$

where  $\gamma$  is the total system throughput,  $S_*$  is the stellar spectrum, and  $M(\lambda, x, y)$  represents the stellar modulation function within the field of view (FoV), including the PSF and speckle patterns. The planetary signal is similarly modulated by the planetary modulation function  $M_p$ , which accounts for both the spatial distribution of the planetary signal and nonideal instrumental effects impacting the planetary spectrum  $S_p$ . Lastly,  $n$  represents fundamental noise sources (photon noise, detector noise, etc.).

Although the planetary modulation function  $M_p$  is introduced here for completeness, its impact is not explicitly considered in Sects. 2 to 4, as its effect on the signal remains relatively minor (<10%) as shown in Sect. 5 (see Fig. 14). This simplification allows for a more straightforward analysis of the dominant noise and systematics affecting detection.

We let denote  $[S]_{LF}$  the low-frequency content of  $S$  (with a certain cutoff resolution,  $R_c$ ) and  $[S]_{HF} = S - [S]_{LF}$  its high-frequency content (see Appendix A). In the following, we consider a cutoff resolution of  $R_c = 100$  (unless otherwise stated), since Bidot et al. (2024) have shown that this is the optimum cutoff resolution in order to suppress speckles efficiently. We will therefore assume that the speckle signature is negligible beyond this cutoff resolution. We define the Hermitian scalar product in the spectral dimension as  $\langle S(\lambda), S'(\lambda) \rangle = \sum_i S^*(\lambda_i) \times S'(\lambda_i)$ . The norm is then given by  $\|S(\lambda)\| = \sqrt{\langle S(\lambda), S(\lambda) \rangle} = \sqrt{\sum_i |S(\lambda_i)|^2}$ .

Initially, the data cube is filtered ( $S_{\text{res}}$ ) by estimating and subtracting low-frequency spectral variations along with the stellar component. Then, molecular mapping consists in applying cross-correlations within the spectral dimension in each spaxel between the residual signal,  $S_{\text{res}}$ , and templates of planets spectra,  $\hat{t}$ , which depend on various properties (temperature, C/O ratio, metallicity, radial velocity, etc.). This process aims to discern the planet's faint signal from the noise: with an appropriate model (and matching radial velocity), a correlation peak at the planet's location can be found. Detailed calculations and mathematical formalism are provided in Appendix B. Thus, at a given radial velocity, the 2D cross-correlation function (CCF) is expressed as

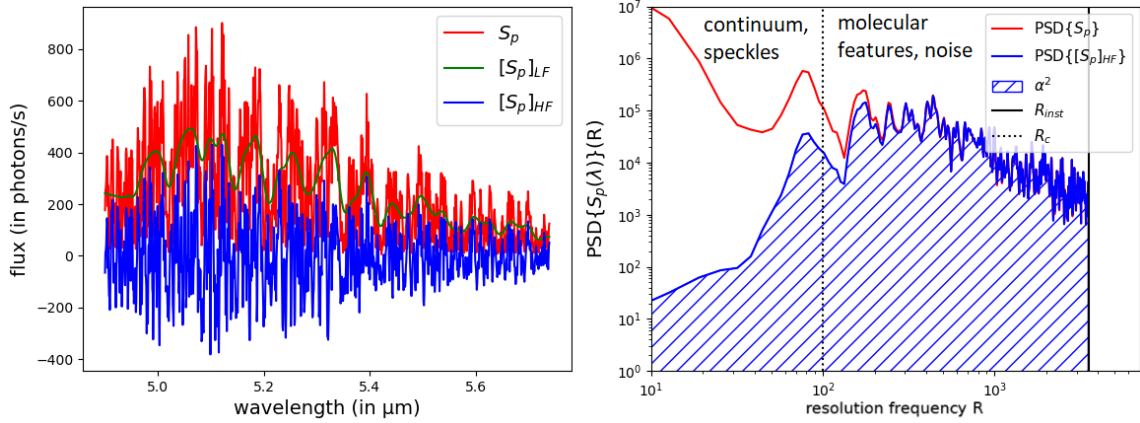
$$\begin{aligned} \text{CCF}(x, y) &= \langle S_{\text{res}}(\lambda, x, y), \hat{t}(\lambda) \rangle \\ &= \underbrace{\alpha(x, y) \cos \theta_p}_{\text{measured signal}} - \beta(x, y) \\ &\quad + \underbrace{\langle n(\lambda, x, y) + n_{\text{sys}}(\lambda, x, y), \hat{t}(\lambda) \rangle}_{\text{noise}} \end{aligned} \quad (2)$$

This is the equation that governs the analysis here. The first term,  $\alpha \cos \theta_p$ , represents the projection of the template onto the observed planetary spectrum, quantifying its spectral content. Where  $\alpha = \|\gamma[M_p S_p]_{HF}\|$  and  $\cos \theta_p$  denotes the similarity factor (correlation) between the template and the observed planetary spectrum (after high-pass filtering), describing the intrinsic mismatch between model and reality. The signal of interest in molecular mapping,  $\alpha$ , is easily predicted and quantified in the Fourier domain as the high-frequency part of the planetary spectrum, lying between the low cutoff frequency and the maximum frequency determined by the instrument's spectral resolution (see Fig. 2)<sup>1</sup>:

$$\alpha(x, y) = \sqrt{\sum_i \text{PSD}\{\gamma(\lambda)[M_p(\lambda, x, y)S_p(\lambda)]_{HF}\}(R_i)}. \quad (3)$$

Fig. 2 illustrates that the higher the cutoff resolution, the lower the spectral richness. This is the purpose of molecular mapping:

<sup>1</sup> This follows from the Parseval-Plancherel theorem, which ensures the conservation of energy in the Fourier domain.



**Fig. 2.** Left: thermal emission spectrum of a planet at 500 K with the Exo-REM atmospheric model (Charnay et al. 2018) downgraded to 1SHORT band resolution ( $R \approx 3500$ ) of MIRI/MRS. Right: PSD of the same spectrum with (blue) and without (red) high-pass filtering. The PSD's limits are defined by the cutoff resolution,  $R_c$ , and the instrumental resolution,  $R_{inst}$ . Due to the Gaussian filter, some spectral content may remain below  $R_c$ , while no lines exist beyond  $R_{inst}$  (see Appendix A).

part of the signal is sacrificed in order to get rid of speckles (or other low-frequency noise). Lastly,  $\alpha$  depends solely on the flux and properties of the planet, particularly its temperature: higher planetary temperatures mean fewer atmospheric lines (i.e.,  $\alpha$  will be smaller for the same flux).

The second term,  $\beta = \|[M_p S_p / S_*]_{LF} \gamma [S_*]_{HF} \cos \theta_*$ , represents the projection of the template onto the lines of the residual stellar spectrum and indicates the amount of signal of interest that is lost.  $\cos \theta_*$  denotes the similarity factor between the template and the observed stellar spectrum. Indeed,  $\beta$  is a self-subtraction term representing the similarity degree between the star and the planet that cannot be used to discriminate the planet from the star (since subtracting the stellar component also eliminates common lines). This term is minimized when the planetary spectrum significantly differs from the stellar spectrum (e.g., when their temperatures are sufficiently different). It is also minimized if the Doppler shift between the planet and the star and/or the instrumental resolution is sufficiently important.

The last term represents the noise projection induced by the correlation. Unlike previous work (Bidot et al. 2024), where it was assumed that all modulations were low-frequency ( $M \approx [M]_{LF}$ ), we assume here that the modulation may contain high-frequency components. As a result, we introduce a new term,  $n_{\text{sys}}(\lambda, x, y) = [M(\lambda, x, y)]_{HF} \gamma(\lambda) S_*(\lambda)^2$ . Instrumental effects, as well as signal extraction and calibration processes that introduce high-frequency artifacts, distort the original spectral information, projecting onto the template and creating a non-uniform CCF signal that effectively acts as a noise source. It is therefore crucial to consider this new term, as we show that it is the primary source of systematic noise. Similar to speckle noise, this systematic noise term is directly proportional to the integrated stellar flux, thereby setting new detection limits.

## 2.2. Signal and noise calculation

First, the S/N of the CCF is simply given by

$$\text{S/N}(x, y) = \frac{E[\text{CCF}(x, y)]}{\sqrt{\text{Var}[\text{CCF}(x, y)]}} = \frac{\alpha(x, y) \cos \theta_p - \beta(x, y)}{\sigma_{\text{CCF}}(x, y)}. \quad (4)$$

<sup>2</sup> If only speckles were present as systematic effects,  $M \approx [M]_{LF}$  and  $n_{\text{sys}}$  would be null. However, some systematic effects introduce high-frequency modulations in the stellar halo, making this term non-zero and generating systematic noise.

To go a step further, a description of the various fundamental noise sources and their propagation in the CCF through correlation is needed:

$$n(\lambda, x, y) = n_{\text{halo}}(\lambda, x, y) + n_{\text{bkgd}}(\lambda) + n_{\text{dc}} + n_{\text{RON}}, \quad (5)$$

where  $n_{\text{halo}}$  is the photon noise associated with the stellar halo,  $n_{\text{bkgd}}$  is the photon noise from the background emission,  $n_{\text{dc}}$  is the detector dark current noise, and  $n_{\text{RON}}$  is the readout noise. The latter three components are assumed to be spatially homogeneous across the FoV.

If there is a spatial dependence of the noise, we assume that its statistics is axisymmetric, meaning that the noise variance only depends spatially on the separation  $\rho = \sqrt{x^2 + y^2}$  from the star. We also assume that the various noise sources are independent of each other, and that there is no spectral covariance for fundamental noises (although spatial covariance is permitted). Deviations from these assumptions could introduce slight errors in the subsequent estimates. Using template's normalization ( $\sum_i \hat{I}^2(\lambda_i) = 1$ ) and the spectral non-covariance assumption for fundamental noises, we can write:

$$\begin{aligned} \sigma_{\text{CCF}}^2(\rho) &= \text{Var}[\langle n(\lambda, x, y), \hat{I}(\lambda) \rangle] + \text{Var}[\langle n_{\text{sys}}(\lambda, x, y), \hat{I}(\lambda) \rangle] \\ &= \text{Var}[\langle n_{\text{halo}}(\lambda, x, y) + n_{\text{bkgd}}(\lambda) + n_{\text{dc}} + n_{\text{RON}}, \hat{I}(\lambda) \rangle] \\ &\quad + \sigma_{\text{sys}}^2(\rho) \\ &= \sum_i \hat{I}^2(\lambda_i) \times (\text{Var}[n_{\text{halo}}(\lambda_i, x, y)] \\ &\quad + \text{Var}[n_{\text{bkgd}}(\lambda_i)] + \text{Var}[n_{\text{dc}}] + \text{Var}[n_{\text{RON}}]) + \sigma_{\text{sys}}^2(\rho) \\ &= \sum_i \hat{I}^2(\lambda_i) \times (\sigma_{\text{halo}}^2(\lambda_i, \rho) + \sigma_{\text{bkgd}}^2(\lambda_i)) \\ &\quad + \sigma_{\text{dc}}^2 + \sigma_{\text{RON}}^2 + \sigma_{\text{sys}}^2(\rho) \end{aligned} \quad (6)$$

with

$$\begin{aligned} \sigma_{\text{sys}}^2(\rho) &= \text{Var}[\langle n_{\text{sys}}(\lambda, x, y), \hat{I}(\lambda) \rangle] \\ &= \text{Var}[\langle [M(\lambda, x, y)]_{HF} \gamma(\lambda) S_*(\lambda), \hat{I}(\lambda) \rangle]. \end{aligned} \quad (7)$$

To simplify the explanation, we note

$$\sigma_{\text{halo}}^2(\rho) = \sum_i \hat{I}^2(\lambda_i) \times \sigma_{\text{halo}}^2(\lambda_i, \rho) \quad (8)$$

and

$$\sigma_{\text{bkgd}}^2 = \sum_i I^2(\lambda_i) \times \sigma_{\text{bkgd}}^2(\lambda_i). \quad (9)$$

$\sigma^2$  then denotes the variances per spectral channel and  $\sigma'^2$  the variances projected into the CCF (no distinction is made for the readout noise and dark current components).

In order to obtain a Poisson statistic for photon noises, we consider all quantities in electron per detector integration time (DIT) and per pixel. In addition, since the planet's signal is spatially distributed across multiple spaxels, we conduct a spatial integration over a box of  $A_{\text{fwhm}}$  spaxels (equivalent to the size of the PSF's full width at half maximum (FWHM)) to optimize the S/N (Ruffio et al. 2019). For  $N_{\text{int}}$  integrations, noise components combine quadratically (assuming independence), while the signal combines linearly. However, as the systematic noise term  $n_{\text{sys}}$  is directly proportional to the integrated stellar flux, its variance scales with  $N_{\text{int}}^2$ . If the data are dithered, the noise statistics will be modified and it is needed to multiply by a corrective factor  $R_{\text{corr}}$  (see Appendix C), otherwise  $R_{\text{corr}} = 1$ . The S/N is thus expressed as

$$S/N(\rho) = \frac{\sqrt{N_{\text{int}}}(\alpha_{\text{fwhm}} \cos \theta_p - \beta_{\text{fwhm}})}{\sqrt{A_{\text{fwhm}} R_{\text{corr}} (\sigma_{\text{halo}}^2(\rho) + \sigma_{\text{bkgd}}^2 + \sigma_{\text{dc}}^2 + \sigma_{\text{RON}}^2 + N_{\text{int}} \sigma_{\text{sys}}^2(\rho))}}, \quad (10)$$

where  $\alpha_{\text{fwhm}}$  and  $\beta_{\text{fwhm}}$  are the same quantities but integrated over the FWHM (so they no longer have any spatial dependency). Since the quantity  $\alpha_{\text{fwhm}} \cos \theta_p - \beta_{\text{fwhm}}$  is proportional to the integrated flux of the planet  $F_p$ ,  $\alpha_0 = (\alpha_{\text{fwhm}} \cos \theta_p - \beta_{\text{fwhm}}) \frac{F_p}{F_*}$  can be defined and corresponds to the signal the planet would have if it had the same integrated flux as the star  $F_*$ . The contrast limit (ensuring  $5\sigma$  detection if the contrast between the planet and the star  $F_p/F_*$  is greater) is then defined by

$$C(\rho) = \frac{5 \sqrt{A_{\text{fwhm}} R_{\text{corr}} (\sigma_{\text{halo}}^2(\rho) + \sigma_{\text{bkgd}}^2 + \sigma_{\text{dc}}^2 + \sigma_{\text{RON}}^2 + N_{\text{int}} \sigma_{\text{sys}}^2(\rho))}}{\sqrt{N_{\text{int}}} \alpha_0}. \quad (11)$$

Finally, for long integration times (i.e.,  $N_{\text{int}}$  large) and/or high stellar fluxes (i.e.,  $\sigma_{\text{sys}}^2$  large), the systematic term will dominate and set the detection limit:

$$S/N(\rho) = \frac{\alpha_{\text{fwhm}} \cos \theta_p - \beta_{\text{fwhm}}}{\sigma'_{\text{sys}}(\rho)} \quad \text{and} \quad C(\rho) = \frac{5\sigma'_{\text{sys}}(\rho)}{\alpha_0}. \quad (12)$$

### 2.3. FastCurves

FastCurves<sup>3</sup> is a numerical and analytical tool initially developed by Bidot et al. (2024) to compute the S/N (Eq. (10)) and the contrast limit (Eq. (11)). In this study, we have extended FastCurves to incorporate systematic effects that impact both the noise and the planetary signal. Specifically, the updated version now models systematic noise through the stellar modulation function  $M$  and accounts for distortions in the planetary signal via the planetary modulation function  $M_p$ . Additionally, FastCurves can now estimate performance when considering PCA in the post-processing of data, as discussed in Sect. 7.

First, we compute the planetary signal component with  $\alpha_{\text{fwhm}} = f_{\text{fwhm}} \|\gamma[M_p \mathbb{S}_p]_{\text{HF}}\|$  where  $f_{\text{fwhm}}$  represents the fraction of flux within the PSF's FWHM (the computation of  $M_p$  is discussed below). The planetary spectrum model  $\mathbb{S}_p$  can be selected

<sup>3</sup> Initial version: <https://github.com/ABidot/FastCurves>  
Updated version: <https://github.com/StevMartos/FastYield>

from various atmospheric models such as BT-Settl (Allard et al. 2012), Exo-REM (Charnay et al. 2018), SONORA (Marley et al. 2021), Morley (Morley et al. 2012) or PICASO (Batalha et al. 2019). In FastCurves, we assume that the observed planetary spectrum  $S_p$  and the template  $\hat{t}$  are identical ( $\cos \theta_p = 1$ ). This implies that no prior assumptions regarding the degree of mismatch are considered, allowing the detection limits to be determined by the instrument specifications rather than our limited understanding of planetary models. In reality, it should be acknowledged that this mismatch could potentially constitute a limitation to molecular mapping (since the  $1 - \cos \theta_p$  fraction of the signal will be lost), and disparities between the template and the observed planetary spectrum are likely to persist.  $\beta_{\text{fwhm}}$  can also be computed by considering a BT-NextGen (Allard et al. 2012) stellar spectrum based on its properties. However, both planetary and stellar spectra need to be degraded to the instrumental resolution, adjusted to their respective magnitudes, and Doppler-shifted to their respective radial velocities to calculate these two quantities. Regarding contrast calculation, a magnitude for the planet is unnecessary (only that of the star), as we renormalize  $\alpha_0$  with the stellar flux.

Next,  $\sigma_{\text{halo}}^2$  (and then  $\sigma_{\text{halo}}^{\prime 2}$  according to Eq. (8)) can be calculated by considering a stellar spectrum and PSF profiles derived from calibration data,  $\sigma_{\text{bkgd}}^2$  (and then  $\sigma_{\text{bkgd}}^{\prime 2}$  according to Eq. (9)) by considering a background model (Rigby et al. 2023),  $\sigma_{\text{dc}}$  and  $\sigma_{\text{RON}}$  by considering detector characteristics (Rieke et al. 2015).

A key advancement in this version of FastCurves is the inclusion of systematic noise estimation. The calculation of  $\sigma_{\text{sys}}^2$  requires an estimation of the high-frequency residual modulations in the stellar halo  $[M]_{\text{HF}}$  (see Eq. (7)) and this can be achieved through two approaches:

- Direct measurement from single-star data with high S/N per spectral channel, where systematic noise dominates over fundamental noises ( $M_{\text{data}}$ ).
- End-to-end simulations (e.g., with MIRISim) to derive a model of the stellar (or planetary) modulation function  $M_{\text{sim}}$  (see Sect. 3.2).

With these estimates, we computed the systematic noise component as

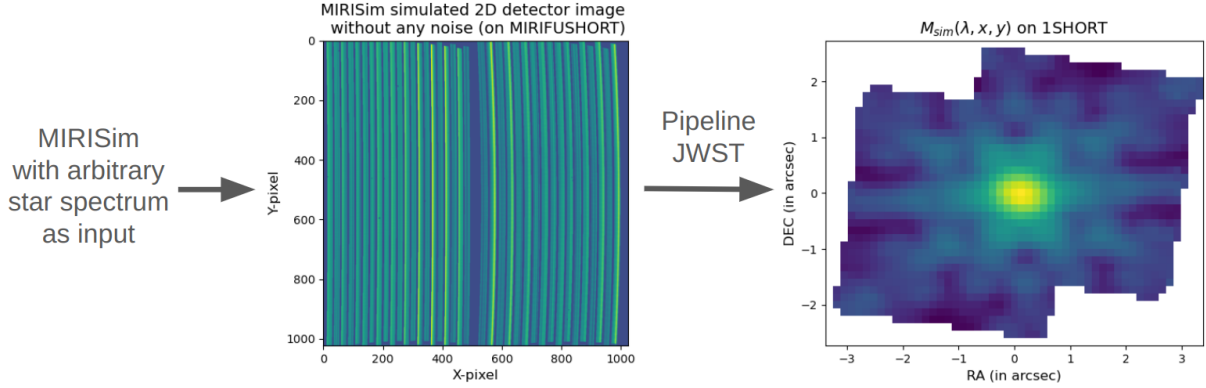
$$\sigma_{\text{sys}}^2(\rho) = \text{Var}[\langle [M_{\text{sim,data}}(\lambda, x, y)]_{\text{HF}} \gamma(\lambda) S_*(\lambda), \hat{t}(\lambda) \rangle]. \quad (13)$$

By directly computing  $\sigma_{\text{sys}}^{\prime 2}$  in this manner, no assumptions regarding spectral covariance are necessary. Incorporating these new features into the framework ensures that both noise and signal distortions are accurately accounted for when estimating molecular mapping detection limits. These enhancements make FastCurves a more robust tool for evaluating the impact of systematics on molecular mapping and optimizing observational strategies for exoplanet detection.

## 3. MIRI/MRS case

### 3.1. Presentation of MIRI/MRS

For this analysis, we selected MIRI/MRS due to its key role in probing exoplanet atmospheres. MIRI/MRS operates in a spectral domain that offers significant advantages. This particular wavelength range is relatively unexplored, presenting numerous opportunities for new discoveries. There is a notable interest in this range, as reflected by the number of studies targeting it (Deming et al. 2024; Henning et al. 2024; Pontoppidan et al. 2024; Miles et al. 2023; Worthen et al. 2024). Furthermore, because MIRI/MRS operates from space, it avoids atmospheric



**Fig. 3.** Left: noiseless MIRI/MRS 2D detector image simulated with MIRISim with an arbitrary stellar spectrum as input. The  $x$ -axis correspond to the spatial position in the FoV and  $y$  axis to the spectral dispersion. Right: cube reconstructed with the JWST pipeline from the noiseless 2D detector image simulated and divided by the input stellar spectrum.

interference, leading to high-quality observations. The instrument operates in a regime where, despite the inherent challenges of large wavelength-to-diameter ratios, it maintains high performance at close angular separations, creating favorable conditions for exoplanet detection and characterization. MIRI/MRS is also particularly convenient to study due to the availability of the MIRISim end-to-end simulator (version 2.4.2; Klaassen et al. 2020) and of the JWST pipeline (version 1.13.4; Labiano-Ortega et al. 2016).

As the only mid-infrared instrument aboard JWST, MIRI’s coverage spans from 4.9 to 27.9  $\mu\text{m}$  (Wells et al. 2015). It offers four operational modes: imaging, coronagraphy, low-resolution spectroscopy (LRS) with  $R \approx 100$ , and medium-resolution spectroscopy (MRS) with  $R \approx 1500\text{--}3500$ . MIRI/MRS functions as an IFS, delivering diffraction-limited spectroscopy. The wavelength range is divided into four spectral channels, each subdivided into three bands denoted as 1SHORT (4.9–5.74  $\mu\text{m}$  at  $R \approx 3515$ ), 1MEDIUM (5.66–6.63  $\mu\text{m}$  at  $R \approx 3470$ ), 1LONG (6.53–7.65  $\mu\text{m}$  at  $R \approx 3355$ ), 2SHORT (7.51–8.77  $\mu\text{m}$  at  $R \approx 3050$ ), etc. (see Table 1 of Argyriou et al. 2023). A single exposure allows one to observe one band of each channel at the same time, meaning that three exposures are required to cover the entire wavelength range. Only the first two channels are considered, since channel 4 has lower sensitivity and exhibits significant photometric errors. Consequently, channel 3, which shares the same detector as channel 4, is also excluded from consideration. By design, the MRS is spatially and spectrally undersampled in channels 1 and 2. For this reason, point source observations are always made using dithered exposures to improve spatial sampling. This involves placing the point source at a different location on the detector on different exposures and combining them by drizzle (Fruchter & Hook 2002; Law et al. 2023). In this way, all MIRI/MRS observations are considered with four-point dithering, providing an effective angular resolution of 0.13 arcsec/pixel for channel 1 and 0.17 arcsec/pixel for channel 2.

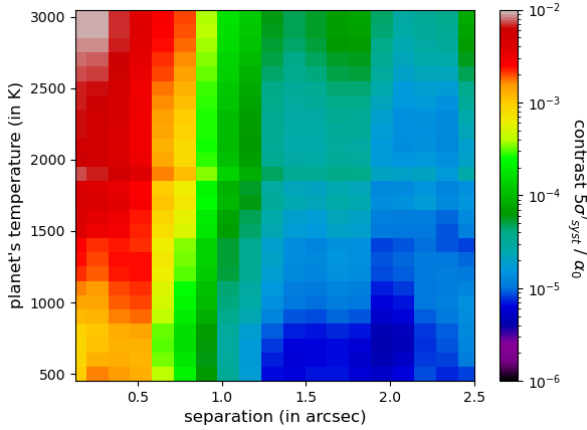
Even in space, Reference Differential Imaging (RDI) is ultimately limited by PSF stability and pointing reproducibility (Ruffio et al. 2024; Málin et al. 2024; Boccaletti et al. 2024). This limitation leaves room for molecular mapping approaches, which remain valuable for both detection and characterization, despite requiring the sacrifice of the continuum. This motivates the need to quantify the detection capabilities of molecular mapping in various observational scenarios, allowing for a discussion of its potential complementarity to RDI.

### 3.2. MIRISim + JWST pipeline as a systematic estimator

The JWST pipeline (Labiano-Ortega et al. 2016) is designed to handle the processing and calibration of raw data acquired from the JWST. MIRISim (Klaassen et al. 2020), an end-to-end simulator, is specifically crafted to produce realistic MIRI data, incorporating various effects such as fringing and distortion. MIRISim outputs are uncalibrated 2D detector images, which are JWST pipeline-compatible inputs. The scene to be observed and the observation parameters can be arbitrarily defined.

Having an end-to-end simulator enables us to estimate systematic effects through the following approach: we generate a 2D detector image using MIRISim without any fundamental noise while injecting an arbitrary stellar spectrum. We then use the JWST pipeline to reconstruct and calibrate a data cube, which we subsequently normalize by the input stellar spectrum to derive an estimate of the stellar modulation function  $M_{\text{sim}}$  (see Fig. 3). Similarly, we can estimate the planetary modulation function  $M_p$  using the same approach as  $M_{\text{sim}}$ , but by injecting a planetary spectrum into MIRISim instead of a stellar spectrum. Although the renormalization step removes the direct influence of the input spectrum, residual dependencies between different injected spectra can still arise due to the nonlinearity of the signal projection by the instrument and its subsequent extraction by the pipeline. These effects become noticeable when strong spectral features are present (particularly in  $M_p$  estimations), introducing minor variations in the derived modulation functions (see Fig. 16). However, these differences remain negligible in practice, making a single simulation sufficient for estimating systematic noise through  $M_{\text{sim}}$ . Additionally, the modulation function is shaped by signal extraction and calibration procedures during cube reconstruction, as well as the object’s position within the FoV. Thus,  $M_{\text{sim}}$  serves as a valuable tool for approximating spatial systematic noise in the CCF using Eq. (13).

Nevertheless, it is crucial to bear in mind that the level of contrast induced by systematic effects represents a detection limit (Eq. (12)). Indeed,  $5\sigma'_{\text{sys}}/\alpha_0$  remains independent of exposure time and the star’s magnitude, given that both  $\sigma'_{\text{sys}}$  and  $\alpha_0$  are directly proportional to the integrated stellar flux. It is therefore possible to calculate the contrast limit for different planet temperatures, showing the detection limits that MIRI/MRS can reach with molecular mapping according to systematics estimated with MIRISim (see Fig. 4). Notably, lower temperatures of the planet yield better contrast, as the spectral content ( $\alpha_0$ )



**Fig. 4.** MIRI/MRS molecular mapping detection limits set by systematics estimated with FastCurves and MIRISim on 1SHORT with BT-Settl models and  $R_c = 100$ .

increases with decreasing temperature, while  $\sigma'_{\text{syst}}$  does not vary significantly with temperature. This raises the important question of whether this systematic noise floor actually appears in the data.

#### 4. Identification of systematics in the noise budget

To validate the accuracy of noise level estimates obtained with FastCurves and MIRISim, we compare our results with both on-sky and simulated data. First of all, we analyze a particular case in detail (Sect. 4.1), highlighting the presence of systematics and validating the fundamental and systematic noise estimates with FastCurves. Next, we extend the validation of fundamental and systematic noise estimates to a diverse set of observational scenarios (Sect. 4.2), demonstrating the reliability of the FastCurves noise model. Finally, in the last two subsections, we show how to improve detection performance through the optimization of cutoff resolution selection (Sect. 4.3), and how to optimize observation time in the presence of systematics (Sect. 4.4).

##### 4.1. Particular case: CT Cha b

To begin with, we consider the on-sky data of CT Cha b with MIRI/MRS, which has been the subject of an observing program for the spectroscopy of its circumplanetary by Rab et al.<sup>4</sup>. Originally discovered and characterized with the IFU VLT/SINFONI by Schmidt et al. (2008), CT Cha b exhibits a magnitude of  $K_s = 14.9$ , an effective temperature of  $T_{\text{eff}} = 2600 \pm 250$  K, and a mass of  $M = 17 \pm 6 M_{\text{jup}}$ , classifying it as a sub-stellar companion (brown dwarf). Positioned at an angular separation of  $\rho \approx 2.5''$  from its host star, we propose a new basic characterization of CT Cha b with molecular mapping in Sect. 6, supplementing previous characterizations proposed post-discovery by Patience et al. (2012), Bonnefoy et al. (2014) and Wu et al. (2015).

On these data cubes, a CCF can be computed using an arbitrary template, as described earlier. This allows us to estimate an empirical spatial standard deviation for each separation  $\sigma_{\text{CCF}}^{\text{spatial}}$ . This standard deviation must include all fundamental noises, but especially systematic noise, if any. This way, we compare the

empirical contrast limit, derived from the spatial standard deviation of the CCF ( $\sigma_{\text{CCF}}^{\text{spatial}}$ ), with the analytical contrast limit from FastCurves, taking systematic noise into account ( $\sigma_{\text{CCF}}^{\text{FC,fund+syst}}$ ). Secondly, the standard deviations of fundamental noises (i.e., Poisson and detector noises) are estimated by the pipeline during ramp fitting, propagated through different pipeline stages, and stored in the ERR extension of the data cubes<sup>5</sup>. While this provides an empirical standard deviation  $\sigma_{\lambda}^{\text{ERR}}$  at each point in the cube, it does not take into account spatial variation and therefore systematic noise. We project the mean standard deviation for each separation  $\sigma_{\lambda}^{\text{ERR}}$  into the CCF  $\sigma_{\text{CCF}}^{\text{ERR}}$  (likewise Eq. (8) or (9)), and we compare it to the analytical contrast limit given by  $\sigma_{\text{CCF}}^{\text{FC,fund}}$  calculated with FastCurves, without taking systematics into account (see Fig. 5).

The template that we used to calculate the CCF is the one yielding the best correlation on 1SHORT (see Sect. 6); that is, a BT-Settl spectrum at 2600 K. On 1SHORT, we observe a favorable agreement between the analytical approach (FastCurves) and the on-sky data (see bottom left of Fig. 6). The analytical noise level with systematics corresponds closely to the empirical noise level derived from spatial variance. Similarly, the analytical noise level without systematics also aligns with the empirical noise level computed from the ERR extension, as expected. The slight discrepancies between the different noise levels may be due to the fact that the simulated noiseless cube used to estimate  $M_{\text{sim}}$  has not been calibrated and reconstructed by the pipeline in the exact same way as the data cube (inducing different systematics). They may also be due to the fact that the noise induced by cosmic rays (or other outliers) is not taken into account and from the different assumptions used.

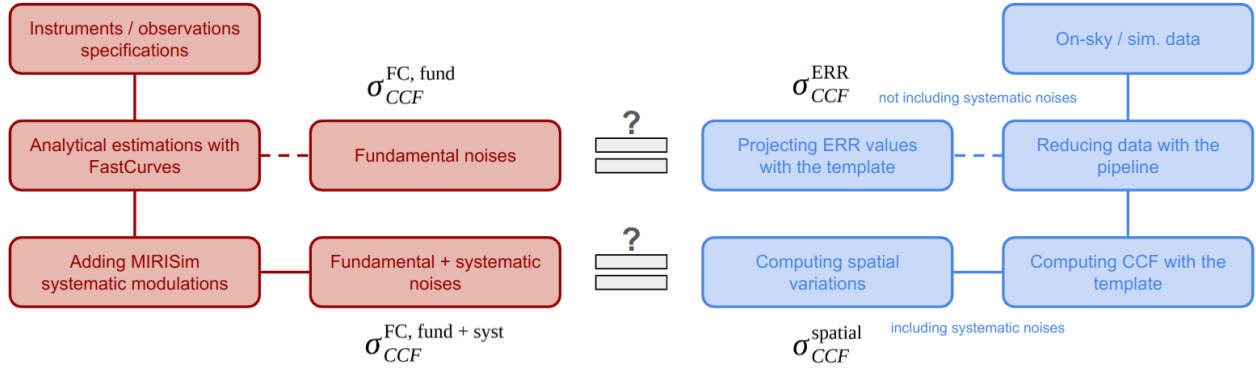
The noise contributions plot (bottom right panel of Fig. 6) indicates that detection is limited by systematic noises at short separations ( $< 0.8''$ ), which is expected due to the high stellar flux in this area. At larger separations, the detection can be limited by other noises if the stellar flux and/or exposure time are not sufficiently high. These two regions are expected to exhibit different statistics: fundamental noises should follow a Gaussian distribution whereas systematic noises may not. This is confirmed through quantile-quantile plots (see Fig. 7), which compare the quantiles of the CCF distribution against the theoretical quantiles of a normal distribution (a Gaussian distribution would approximately lie on the identity line  $y = x$ ). This implies that  $5\sigma$  detection may not carry the same significance in regions dominated by systematic noises (Garvin et al. 2024). Moreover, these Q-Q plots serve as additional validation of the predictions made by FastCurves.

##### 4.2. Validation of the noise model on various stellar cases

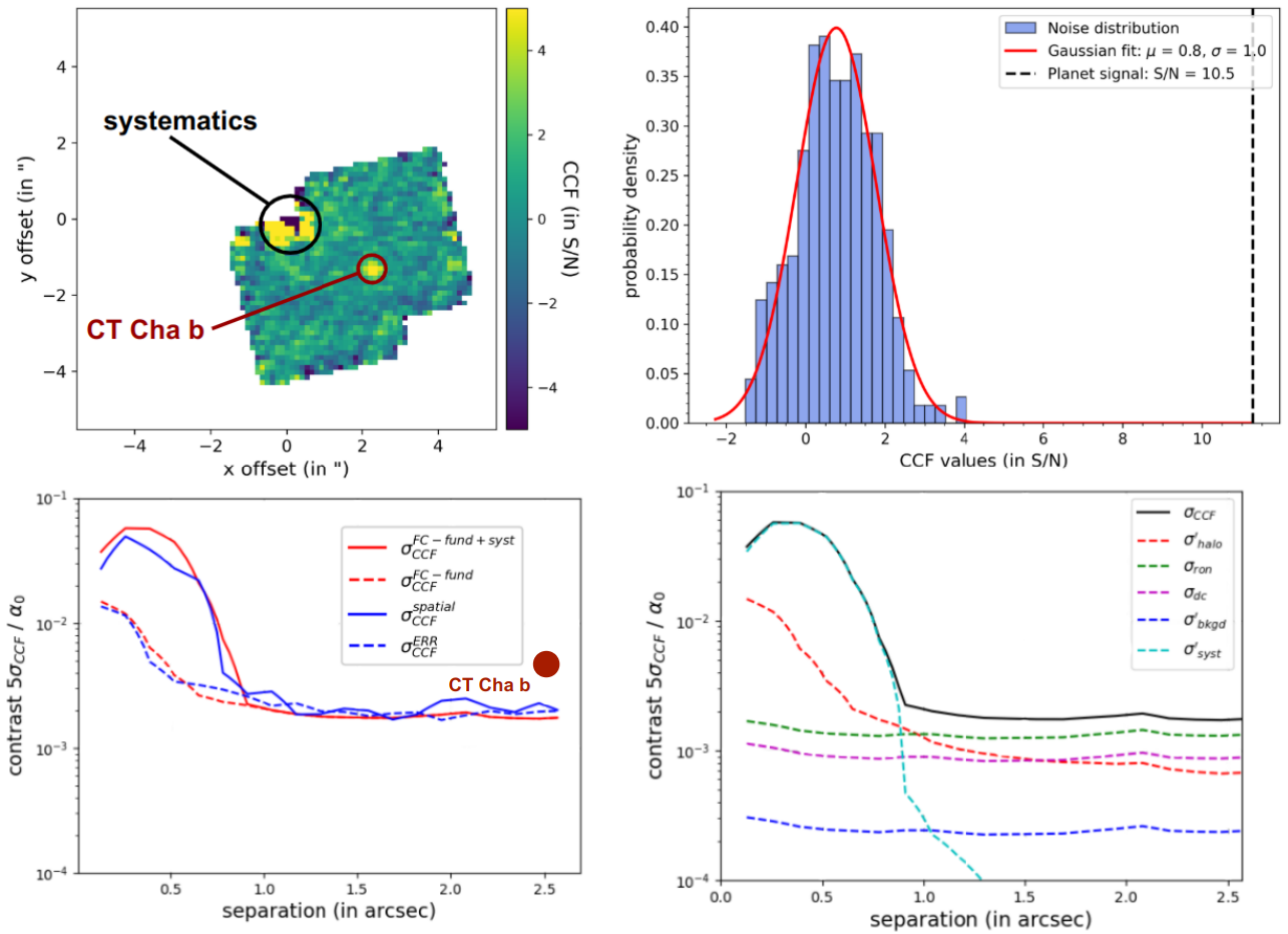
It is essential to verify whether FastCurves accurately estimates noise levels across various observational cases and spectral bands. To achieve this, we compare the analytical contrasts (with and without systematics) with empirical contrasts derived from both spatial variance and ERR extension for each separation, band, and various simulated and on-sky data cases. We observe a high level of agreement between the FastCurves estimates and the simulated data used in Mâlin et al. (2023) across all bands (see Fig. 8). This alignment is expected for systematic noise, given its estimation from MIRISim simulated data as well.

<sup>5</sup> The error values in the ERR extensions may be incorrectly estimated by the pipeline, sometimes differing by factors as large as 10–50. Therefore, the ERR values should be taken with caution.

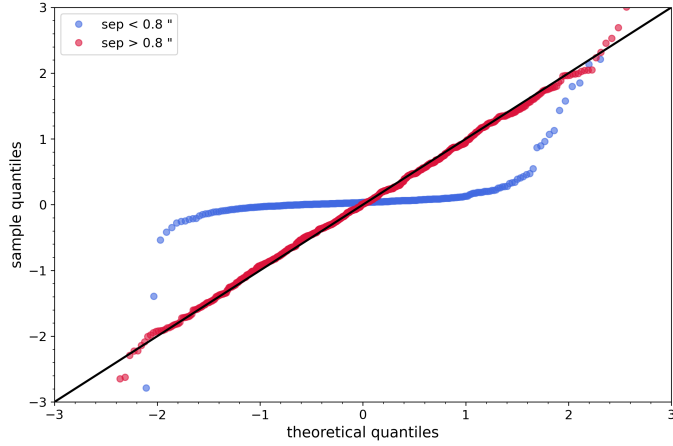
<sup>4</sup> PI: C. Rab, <https://www.stsci.edu/jwst/phase2-public/1958.pdf>



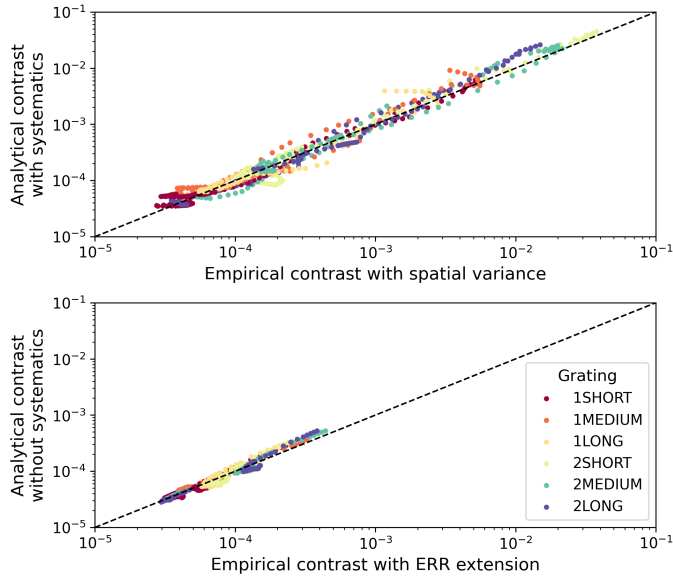
**Fig. 5.** Methodology for comparing analytical (red) and empirical (blue) noise estimates. Empirical errors are obtained from on-sky data cubes by projecting the ERR extension values onto the template, as well as by calculating the spatial standard deviation of the CCF. These empirical values are then compared to the analytical errors derived from FastCurves, both with and without consideration of the systematic noise component (estimated with MIRISim), in addition to the fundamental noise sources.



**Fig. 6.** Top left: CCF of CT Cha b on-sky data on 1SHORT, calculated with a BT-Settl template at 2600 K Doppler shifted with 13.5 km/s and  $R_c = 100$  (the S/N color scale is cropped to  $\pm 5$ ). Top right: histogram of the CCF. Bottom left: contrast curves calculated analytically using FastCurves are shown in red, assuming  $t_{exp} = 56$  min,  $K_* = 8.7$ , and the same planetary spectrum used to compute the CCF. The solid red curve represents the result including both fundamental and systematic noise ( $\sigma_{CCF}^{FC, fund + syst}$ ), while the dashed red curve represents only fundamental noise ( $\sigma_{CCF}^{FC, fund}$ ). Empirical contrast curves derived from on-sky data are shown in blue, where the solid blue curve corresponds to the spatial standard deviation of the CCF ( $\sigma_{CCF}^{spatial}$ ), and the dashed blue curve corresponds to the errors estimated from the ERR extension values ( $\sigma_{CCF}^{ERR}$ ). Bottom right: analytical contrast contributions.



**Fig. 7.** Q-Q plots of the CCF of CT Cha b data on 1SHORT for a separation greater (in red) or lesser (in blue) than  $0.8''$ .

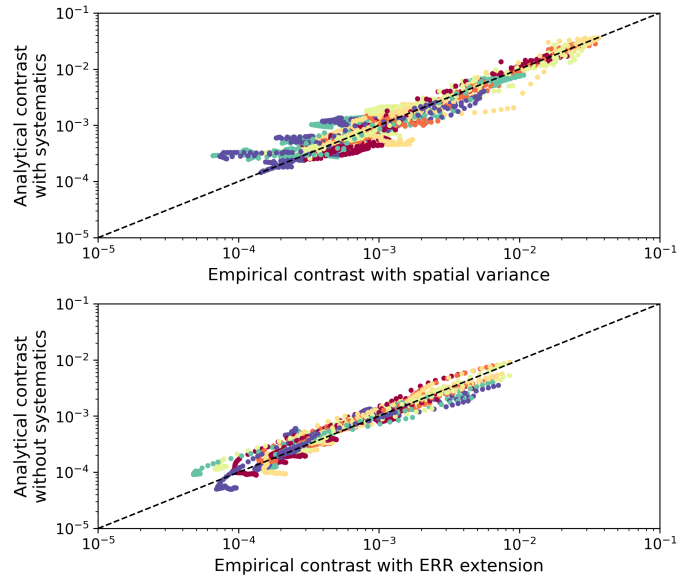


**Fig. 8.** Comparison of analytical noise levels (y axis) with systematics  $\sigma_{\text{CCF}}^{\text{FC,fund+syst}}$  (top) and without systematics  $\sigma_{\text{CCF}}^{\text{FC,fund}}$  (bottom), against empirical noise levels in simulated data (x-axis) from spatial standard deviations  $\sigma_{\text{CCF}}^{\text{spatial}}$  (top) and ERR extensions  $\sigma_{\text{CCF}}^{\text{ERR}}$  (bottom). Each point is a given separation, for a given band and data set. The simulated data considered are those used in [Málin et al. \(2023\)](#) for HR8799 (with  $K_* = 5.24$  and  $t_{\text{exp}} = 171$  min per band) and  $\beta$  Pictoris (with  $K_* = 3.48$  and  $t_{\text{exp}} = 93$  min per band).

In order to have a full comparison of the noise levels estimated with FastCurves and on-sky data, we consider the cases listed in Table 1. Although the agreement between noise levels estimated with FastCurves and on-sky data is relatively good, it is not as good as with simulated data (since the dispersion is wider, see Fig. 9). For noise levels including systematics, it suggests that the systematics simulated by MIRISim may not perfectly mirror those encountered in on-sky observations (as detailed in Sect. 5). Moreover, it seems that in all cases the behavior of fundamental noises is correctly estimated. In conclusion, it can be seen that the noise model used in FastCurves seems to be appropriate and sufficient to recover the noise statistics in the data.

**Table 1.** Considered targets, including stellar magnitude, exposure time, program number, and principal investigator (PI) name.

Target	$K$	$t_{\text{exp}}$ (min)	Program	PI
HD 159222	5.00	18	1050	B. Vandenbussche
del UMi	4.26	7	1524	D. Law
HD2811	7.06	24	6604	K. Gordon
V* GO Tau	9.33	39	1640	A. Banzatti
V* TW Cha	8.62	40	1549	K. Pontoppidan
HD 163296	4.78	7	2025	K. Oberg

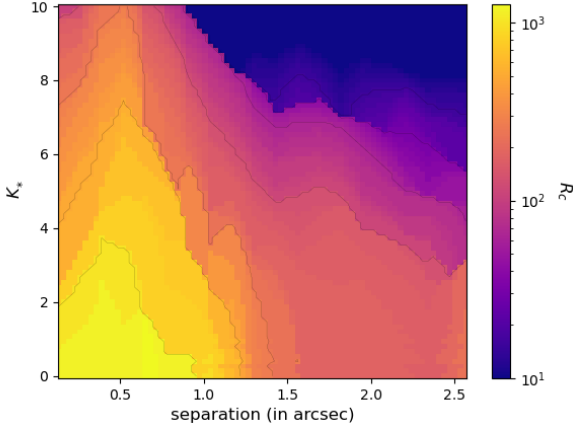


**Fig. 9.** Same as Fig. 8 but with on-sky data listed in Table 1. To recover the agreement between FastCurves and on-sky data for the top panel, systematics (i.e.,  $M_{\text{sim}}$ ) for which fringes and straylight had been incorrectly subtracted by the pipeline had to be considered (see Sect. 5).

Consequently, the initial challenge remains: although the combination of high-contrast imaging and high-resolution spectroscopy with molecular mapping effectively removes speckles that limit detection at short separations, another type of systematic noise continues to constrain detection in this region. One potential strategy to mitigate this systematic noise is by opting for a higher cutoff resolution, but a trade-off must be found to avoid filtering out too much signal (see Sect. 4.3). However, we shall also demonstrate later (see Sect. 7) that a more efficient method of reducing systematic noise is to apply a PCA.

#### 4.3. Impact of systematics on optimal cutoff resolution

To determine the optimum cutoff resolution for each specific case, we calculated the S/N according to Eq. (10) for different cutoff resolutions and star magnitudes. Consequently, we could determine the optimum cutoff resolution (yielding the best S/N) as a function of the star's magnitude and separation (see Fig. 10). This map delineates the most favorable trade-off in cutoff resolution between signal retention and systematic noise suppression. It reveals that the optimum cutoff resolution decreases with increasing separation, reflecting the decrease in systematics attributable to the declining stellar flux. Additionally, for brighter stars, a higher optimal cutoff resolution is



**Fig. 10.** Optimum cutoff resolution as a function of stellar magnitude with a BT-Settl planetary spectrum at 1000 K and a total exposure time of one hour.

preferable for each separation. While estimating the optimal cut-off resolution may not be inherently useful, given that it can be empirically determined when applying molecular mapping to the data, it is noteworthy that in scenarios dominated by systematics, a higher cutoff resolution is required compared to the one needed for eliminating speckles ( $R_c \approx 100$ ).

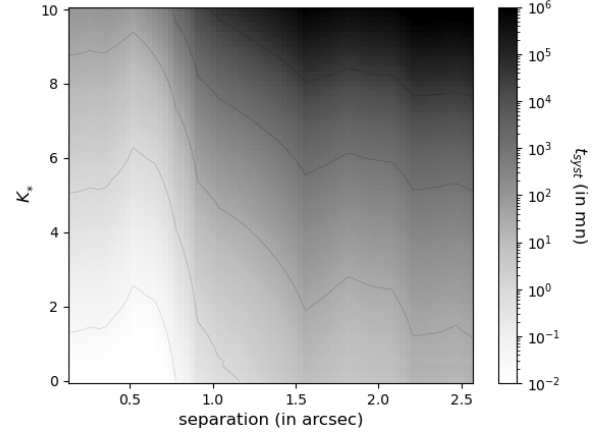
#### 4.4. Impact of systematics on total exposure time

Selecting an appropriate exposure time is crucial, as the S/N does not increase indefinitely with the square root of exposure time in the presence of systematics. Molecular mapping is employed to remove speckles, assuming they are the primary source of systematic noise. Beyond a cutoff resolution of  $R_c = 100$ , speckles become negligible, implying that no residual systematic noise should remain. The key question is to determine the exposure time at which this assumption no longer holds, as additional systematics start to dominate and degrade detection performance. Identifying this threshold helps establish the exposure time limit beyond which molecular mapping alone becomes less effective.

To quantify this limit, we defined the exposure time,  $t_{\text{syst}}$ , at which systematic noise becomes comparable to fundamental noise, marking the point where systematics become predominant:

$$\begin{aligned}
 \sigma_{\text{syst}}^{\text{total}}(\rho) &= \sigma_{\text{fund}}^{\text{total}}(\rho) \\
 \rightarrow N_{\text{int}}^2 \sigma_{\text{syst}}^2(\rho) &= N_{\text{int}}(\sigma_{\text{halo}}^2(\rho) + \sigma_{\text{bkgd}}^2 + \sigma_{\text{dc}}^2 + \sigma_{\text{RON}}^2) \\
 \rightarrow N_{\text{int}} &= \frac{\sigma_{\text{halo}}^2(\rho) + \sigma_{\text{bkgd}}^2 + \sigma_{\text{dc}}^2 + \sigma_{\text{RON}}^2}{\sigma_{\text{syst}}^2(\rho)} \\
 \rightarrow t_{\text{syst}}(\rho) &= \text{DIT} \frac{\sigma_{\text{halo}}^2(\rho) + \sigma_{\text{bkgd}}^2 + \sigma_{\text{dc}}^2 + \sigma_{\text{RON}}^2}{\sigma_{\text{syst}}^2(\rho)},
 \end{aligned} \tag{14}$$

where DIT is the exposure time per integration.  $t_{\text{syst}}$  represents the exposure time at which the S/N begins to reach the detection limit (Eq. (12)), indicating the point at which further observation becomes less beneficial. Figure 11 highlights that the time required to reach the detection limit imposed by systematics increases with a lower stellar flux (due to greater magnitude and/or separation). While this outcome is not surprising, it underscores the feasibility of estimating the optimal



**Fig. 11.** Optimum exposure time as a function of the stellar magnitude on 1SHORT with  $R_c = 100$ .

exposure time for a given observational case and highlights the utility of FastCurves as an exposure time calculator, particularly for optimizing telescope time allocation, in addition to its performance estimation capabilities.

## 5. Contributors to systematics

Since we have highlighted the dominant role of systematics in limiting molecular mapping detection capabilities at short separations for bright targets, we further explore in this section their origin, behavior, and impact. As was previously stated, these systematics (Eq. (13)) arise from any instrument-based imperfections in the signal that include high-frequency content and subsequently project onto the template in the CCF.

We begin this section by listing and describing the potential effects contributing to the observed systematics (Sect. 5.1). Next, we discuss the impact of different effects and correction procedures on the structure and spatial variations in modulations (Sect. 5.2). We then explore how these spatial variations affect noise statistics (Sect. 5.3) and examine how the modulation structure influences the signal (Sect. 5.4).

### 5.1. Possible sources of systematic effects

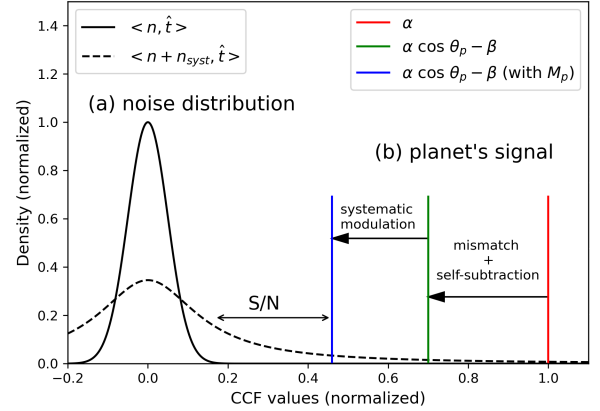
Errors in photometric and wavelength calibration (due to errors in the calibration data, imperfections in re-interpolation operations, cube reconstruction, slice processing, etc.) are likely to induce residuals proportional to the initial flux. A detailed study of the various known effects and in-flight performances of MIRI/MRS is given in Argyriou et al. (2023). Here is a summary of the ones that may impact our study:

- Wavelength calibration: an absolute error in the wavelength solution would not impact detection with molecular mapping, but would simply induce an error in the radial velocity retrieval. If there are relative errors between spaxels for the wavelength solution, this can be more troublesome. In particular, since the planet's signal is integrated over several spaxels, any wavelength shift in the spectra of the different spaxels would result in spectral line broadening, thereby diminishing correlation (see Sect. 5.4). Such shifts could also introduce high-frequency modulations in the stellar halo spectrum, potentially contributing to systematic noise (see Sect. 5.3). The current solution (FLT-6) provides calibration accuracy of a few kilometers per second in channels 1 and 2.

- Photometric calibration: similar to wavelength calibration, an absolute error in photometry would not impact the actual S/N found. However, if photometric calibration errors introduce high-frequency modulations, this could result in diminished correlation (see Sect. 5.4) and systematic noise (see Sect. 5.3). All the effects listed below are likely to induce multiplicative errors (modulations) on the flux of each spaxel, and are therefore considered as photometric systematics.
- Straylight (detector internal scattering): scattered light in MIRI/MRS arises from the scattering of photons within the detector substrate and from light diffraction at the narrow gaps between pixels. This results in a broader PSF and an overlap between the detector’s diffraction pattern and the instrument’s PSF, inducing horizontal band and small spike features in the 3D reconstructed cube (see Fig. 4 of Argyriou et al. 2023). Those features may be either over- or under-subtracted at the percent level by the pipeline.
- Fringing: similar to many infrared spectrometers, MIRI/MRS exhibits notable spectral fringing attributed to Fabry-Perot interference within the detectors (Argyriou et al. 2020). The amplitude (10–30% of the flux) and phase of the fringes depend on the source extension (whether the source is point-like, non-resolved, or extended, resolved) and wavelength. Presently, the JWST pipeline incorporates two dedicated steps for fringe removal. The first correction involves dividing the 2D detector images by a static 2D flat of fringes derived from spatially extended sources, effectively eliminating fringing from the spectra of such sources. But for point sources, this static correction unavoidably leaves residual fringes, as they generate distinct fringe patterns on the detectors compared to those accounted for by the flat. For this reason, the second step is a residual fringe elimination, which iteratively finds and removes the remaining periodic features in the spectrum, reducing the fringe contrast down to less than 6%. Nonetheless, a key question regarding this second correction step remains open for point sources with high spectral richness: either the lines can be identified as fringes by the correction algorithm (and are suppressed), or they can prevent the algorithm from identifying fringes (preventing their suppression; Gasman et al. 2023). Consequently, the latter step is not considered here.
- Resampling noise: this effect is due to an interplay between the cube reconstruction algorithm (re-interpolation of the native detector pixel data into a regular cube grid), spatial undersampling and the curvature of the spectral traces on the detector (Smith et al. 2007). The resampling of unresolved point sources results in aliasing artifacts, manifesting in the data as relatively low-frequency sinusoidal modulations (see Fig. 9 of Law et al. 2023). The amplitude of these artifacts is estimated at around 10% for channels 1 and 2 with four-point dithering (see Fig. 10 of Law et al. 2023).

## 5.2. Structure and spatial variations in the modulations

All of the effects outlined in Sect. 5.1, along with potentially unidentified factors, contribute to high-frequency modulations in the data, resulting in systematic noise and a reduction in correlation strength. On the one hand, if the residual high-frequency modulations of the stellar halo,  $[M]_{\text{HF}}$ , vary spatially, the projection values onto the template in the CCF will also exhibit spatial variations, broadening the noise distribution.



**Fig. 12.** Schematic diagram of the effect of systematics on noise statistics (a) and on the planet’s signal (b). It is important to note that, as previously mentioned, the noise distribution in the presence of systematics will not be Gaussian, making the S/N ratio an unsuitable metric (Garvin et al. 2024).

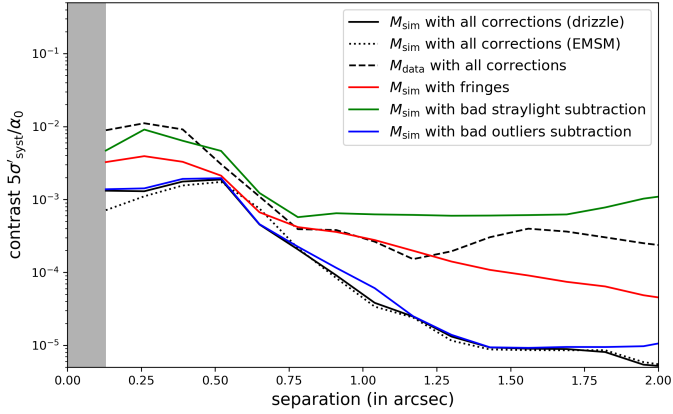
Conversely, if these modulations were spatially uniform, they would not introduce systematic noise. On the other hand, the planetary modulation function,  $M_p$ , describes how the planetary signal is projected by the instrument and subsequently extracted by the pipeline. These nonideal modulations degrade the alignment with the template, thereby reducing correlation effectiveness. In summary, the spatial variations in the stellar modulation function,  $M$ , impact the noise statistics (i.e., the noise distribution (a) in Fig. 12), while the structure of the planetary modulation function  $M_p$  over the planet’s FWHM impacts the correlation strength (i.e., the planet’s signal (b) in Fig. 12).

In Appendix D, we analyze the impact of different instrumental corrections (fringe, straylight, outlier subtraction) and reconstruction methods (drizzle vs. EMSM) on the simulated modulation function  $M_{\text{sim}}$ . The structure and spatial variations in the resulting high-frequency modulations are evaluated using PSD calculations. Our results indicate that, even after applying all pipeline corrections, residual modulations persist, primarily due to aliasing artifacts. Furthermore, uncorrected fringes and straylight significantly affect the PSD of the modulations. We also observe discrepancies between the modulations in simulated and real data, with the latter showing additional modulations likely attributed to uncorrected fringe and straylight effects.

## 5.3. Systematic effects on noise

To assess the impact of different effects on modulation-induced noise (i.e., on the noise distribution (a) in Fig. 12), we calculated systematic noise levels using Eq. (13) with an arbitrary template. We depict the corresponding contrasts in Fig. 13. The results show that straylight is the most critical effect requiring accurate correction, followed by fringing. Poor outlier subtraction or different reconstruction methods appear to have minimal impact. The contrast derived from the observed modulations in on-sky data falls between the levels associated with bad straylight and fringing correction, suggesting that both effects currently limit detection performance in on-sky data<sup>6</sup>.

<sup>6</sup> This result aligns with the necessity of accounting for bad subtraction of straylight and fringes in the estimation of systematics with MIRISim to achieve the agreement seen in Fig. 9 between FastCurves and the on-sky data.



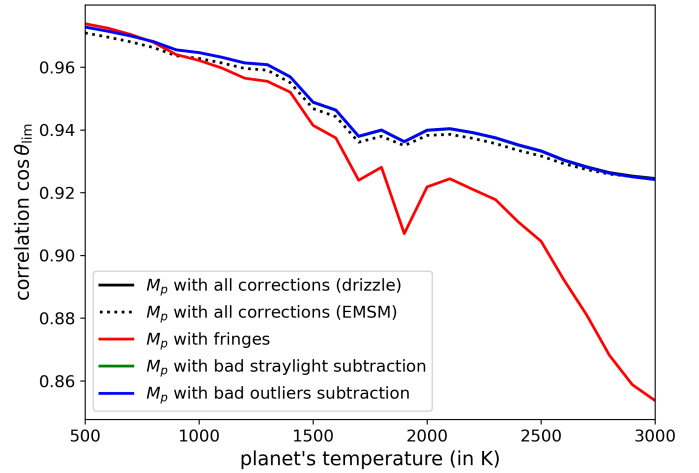
**Fig. 13.** Contrast set by systematics according to various pipeline corrections on 1SHORT with an arbitrary BT-Settl template at 1000 K and  $R_c = 100$ .

A question that remains when straylight and fringes are properly corrected is what the reason is for the remaining contrast level (solid black line in Fig. 13). In line with the previous discussion, it seems reasonable to speculate that this residual contrast stems from residual aliasing artifacts post high-pass filtering. To test this hypothesis, we simulated photometric calibration errors of the same nature as resampling noise (see Appendix E.1). We find that when the simulated systematics have the same amplitude as that measured for the aliasing artifact in the data ( $\approx 7\%$ ), we recover the order of magnitude of the contrast limit obtained with  $M_{\text{sim}}$  (with all corrections, i.e., solid black line in Fig. 13). This suggests that resampling noise is likely the primary systematic effect when straylight and fringes are properly corrected.

At the same time, we simulated the impact of wavelength calibration errors on contrast (see Appendix E.2). In terms of magnitude, the effects of wavelength calibration errors (of a few kilometers per second) on noise appear to be largely negligible. In fact, a calibration error of 100–1000 km/s would be required to account for this systematic noise floor (solid black line in Fig. 13).

In conclusion, the systematic noise level appears to be primarily influenced by straylight, fringing, and resampling noise. To alleviate the contribution of the first two, corrective measures are essential during both the correction stages and the subsequent signal extraction and cube reconstruction processes, as mishandling at any stage can exacerbate these effects (Gasman et al. 2024).

Secondly, despite the benefits of dithering in reducing aliasing artifacts through improved spatial sampling, their presence remains significant. To mitigate the resampling noise effect (mentioned in Sect. 5.1), it may be wise to avoid the interpolation steps that are introduced in the cube reconstruction step. One potential approach involves directly using 2D calibrated detector images and implementing an alternative molecular mapping technique, similar to the forward modeling method proposed in Ruffio et al. (2024) for JWST/NIRSpec/IFU. This approach also benefits from the use of spline filtering as a high-pass filter, which allows bad pixels to be ignored, unlike other convolution filters that propagate interpolation errors from bad pixels. Although this technique has shown great improvement for high-contrast spectroscopy with NIRSpec/IFU, adapting it to MIRI/MRS presents additional challenges, particularly regarding fringing, which must first be addressed appropriately.



**Fig. 14.** MIRI/MRS correlation limits set by MIRISim systematics according to various pipeline corrections on 1SHORT with  $R_c = 100$ . Blue, green, and black curves are overlaid.

This forward modeling approach is currently under investigation (Bidot et al. in prep.), incorporating recent calibration improvements for point source reduction, as described in Argyriou et al. (2020); Gasman et al. (2023); Gasman et al. (2024).

#### 5.4. Systematic effects on correlation

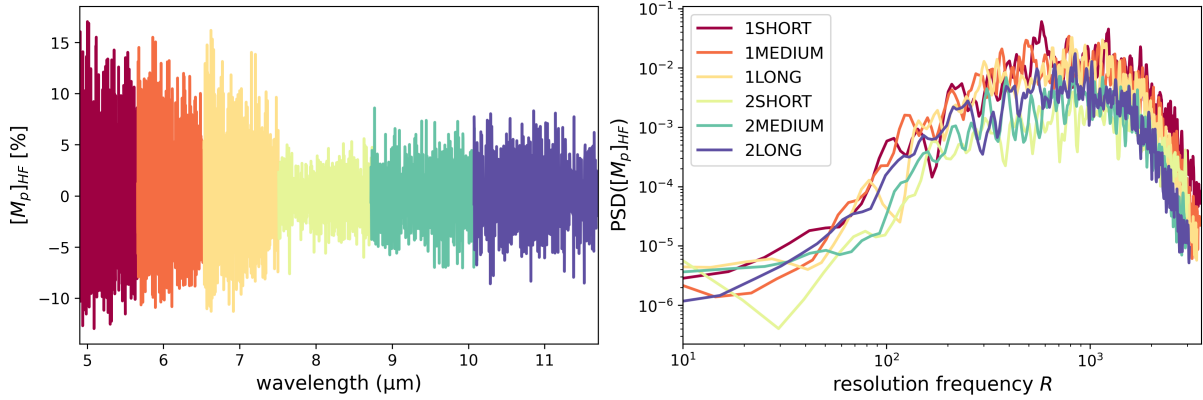
To assess the influence of systematics on correlation (i.e., the planet's signal (b) in Fig. 12), we reiterated the method previously employed: we simulated noiseless cubes using MIRISim and processed them through the pipeline by injecting different planetary spectra (at different temperatures). From these cubes, we integrated the planetary flux over the FWHM and estimated the drop in correlation according to

$$\cos \theta_{\text{lim}} = \frac{\langle \gamma(\lambda)[M_p(\lambda)\mathbb{S}_p(\lambda)]_{\text{HF}}, \gamma(\lambda)[\mathbb{S}_p(\lambda)]_{\text{HF}} \rangle}{\|\gamma(\lambda)[M_p(\lambda)\mathbb{S}_p(\lambda)]_{\text{HF}}\| \times \|\gamma(\lambda)[\mathbb{S}_p(\lambda)]_{\text{HF}}\|}, \quad (15)$$

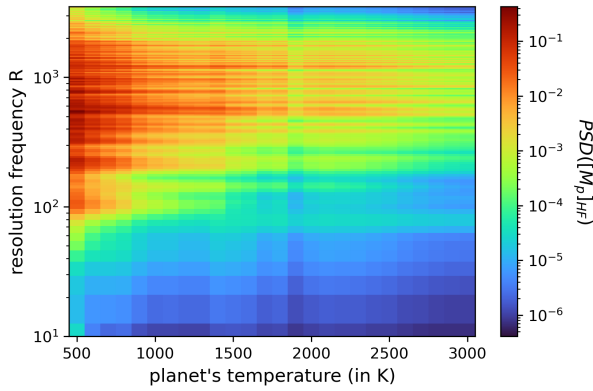
where  $\mathbb{S}_p(\lambda)$  is the injected planetary spectrum model. This means that the reduction in correlation arises not only from spectral mismatches between the observed spectrum and the template used but also from systematic effects, as quantified here. Thus, we assume that the effective correlation loss from both effects is simply given by the product  $\cos \theta_{\text{lim}} \times \cos \theta_p^7$ .  $\cos \theta_{\text{lim}}$  represents the maximal correlation achievable given the systematic distortions (see Fig. 14). We find that the most important effect to correct properly for the correlation is the fringing, with a negligible impact observed from other effects. Also, the greater the temperature, the more the correlation is diminished by the modulations. This is because the lower the temperature, the greater the spectral content and the more it will dominate high-frequency systematic modulations. Finally, the same question arises again of why, when the fringes are perfectly corrected, there is a remaining drop in correlation (solid black line in Fig. 14).

Similarly to the approach described earlier, we estimated the reduction in correlation caused by simulated photometric effects akin to resampling noise (see Appendix F.1). This revealed that the systematic decline in correlation observed in the data cannot be attributed to resampling noise, given that a photometric

<sup>7</sup> This assumption holds strictly when the systematic deviations of the observed planetary spectrum are entirely uncorrelated with the template used.



**Fig. 15.** Left: overall high-filtered systematic modulations estimation while injecting a BT-Settl spectrum at 1000K in MIRISim with  $R_c = 100$ . Right: PSDs of these modulations.



**Fig. 16.** PSDs of high-filtered modulations of the planetary spectrum as a function of its temperature on 1SHORT with  $R_c = 100$  (same as the right panel of Fig. 15 but for different temperatures).

error of 20–50% would be required to explain it. This is partly due to the fact that the modulations of the aliasing artifacts are relatively low-frequency compared to the spectral characteristics of the planetary spectra.

In the same way, it is possible to estimate the reduction in correlation caused by simulated wavelength calibration errors (see Appendix F.2). It suggests that the expected wavelength calibration errors also do not appear to be the dominant source of the systematic drop in correlation (given that it would take an error of around 60 km/s to explain it).

In summary, the observed drop in correlation when fringes are corrected does not appear to be due to resampling noise or wavelength calibration errors. Instead, this reduction in correlation is attributed to very high-frequency ( $R \approx 1000$ ) modulations common to all spaxels, rather than relative modulations between spaxels (see Fig. 15). This overall modulation is strongly influenced by the injected spectral content: the greater the content (i.e., the lower the temperature), the more pronounced the modulations (see Fig. 16)<sup>8</sup>. The origin of this overall modulation remains unclear, but it is reasonable to speculate that it arises

<sup>8</sup> This explains why this modulation dominates over aliasing artifacts for the planetary modulation function over the FWHM, whereas it is the opposite for the stellar modulation function (since this modulation is not clearly visible in the left panel of Fig. D.1). Even if this modulation dominated the stellar modulation function, it would not introduce noise, as it is spatially homogeneous.

from common photometric calibration errors and/or interpolation procedures. Finally, given that this systematic effect operates at high frequency, increasing the cutoff resolution would not improve correlation. However, even in the worst-case scenario – at high temperatures and with poor fringe calibration – the signal lost due to systematics remains below 15%. This implies that systematics will mainly limit detections by impacting spatial noise statistics, meaning they have a greater impact on (a) than on (b) in Fig. 12.

## 6. Signal and correlation estimations

While FastCurves has proven reliable for estimating noise levels, its accuracy in modeling the signal itself remains to be assessed. To address this, we again considered the case of CT Cha b. First, we estimated the planetary spectrum model that best correlates with the observed data spectrum using

$$\cos \theta_{\text{est}} = \frac{\langle \hat{d}(\lambda), \hat{i}(\lambda) \rangle}{\|\hat{d}(\lambda)\|}, \quad (16)$$

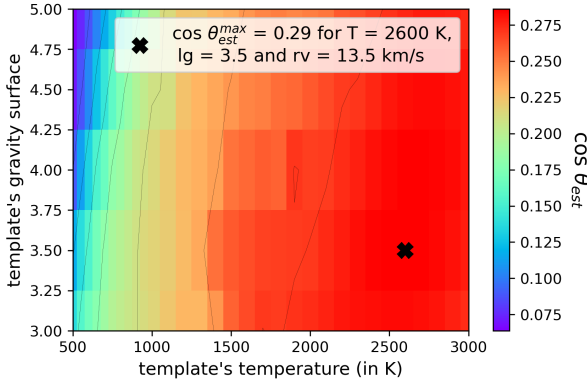
where  $\hat{d}$  is the signal of the planet in the high-filtered cube,  $S_{\text{res}}$ , integrated over the FWHM. The spectrum that appears to be most similar to the CT Cha b spectrum is a BT-Settl spectrum at 2600 K with a surface gravity of 3.5 and a Doppler shift of around 13.5 km/s (achieving a correlation of around 0.29, see Fig. 17)<sup>9</sup>.

Now, the idea is to deduce the intrinsic mismatch,  $\cos \theta_p$ , from the estimated correlation,  $\cos \theta_{\text{est}}$ . Neglecting self-subtraction, it is possible to write (using the fact that  $\langle n + n_{\text{sys}}, \hat{i} \rangle \approx 0$  on average):

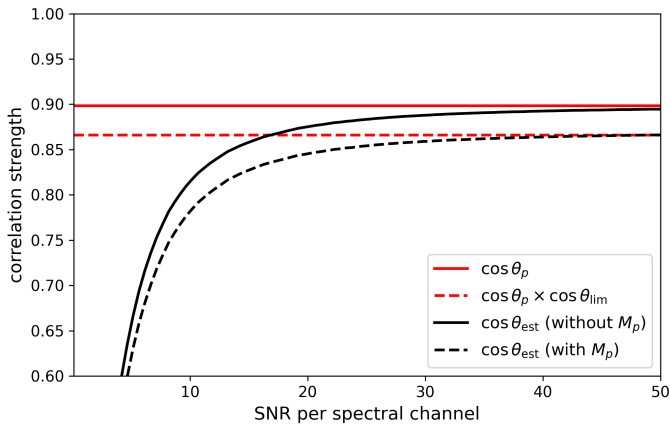
$$\begin{aligned} \hat{d}(\lambda) &= \gamma(\lambda)[M_p(\lambda)S_p(\lambda)]_{\text{HF}} + n(\lambda) + n_{\text{sys}}(\lambda) \\ \rightarrow \cos \theta_{\text{est}} &= \frac{\alpha}{\|\hat{d}(\lambda)\|} \times \cos \theta_{\text{lim}} \times \cos \theta_p \\ \rightarrow \underbrace{\cos \theta_{\text{est}}}_{\text{estimated correlation}} &= \underbrace{\cos \theta_n}_{\text{noise-induced correlation loss}} \times \underbrace{\cos \theta_{\text{lim}}}_{\text{systematic-induced correlation loss}} \times \underbrace{\cos \theta_p}_{\text{model-induced correlation loss}}, \end{aligned} \quad (17)$$

defining  $\cos \theta_n = \alpha/\|\hat{d}\|$ , representing the drop in the estimated correlation due to fundamental noises. Therefore, the estimated correlation ( $\cos \theta_{\text{est}}$ ) may not accurately reflect the intrinsic

<sup>9</sup> Minimizing a  $\chi^2$  in the same way as in Mâlin et al. (2023) yields to the same parameters.



**Fig. 17.** Estimated correlation between CT Cha b high-filtered data spectrum and BT-Settl spectra on 1SHORT with  $R_c = 100$ .



**Fig. 18.** Impact of noise on the estimation of the correlation on 1SHORT with  $T_p = 1000$  K and  $R_c = 100$ . A BT-Settl spectrum is associated with the observed spectrum and a SONORA spectrum with the template to create an artificial mismatch ( $\cos \theta_p \approx 0.9$ ). For each S/N per spectral channel, Gaussian noise is added to the observed spectrum to simulate the noise.

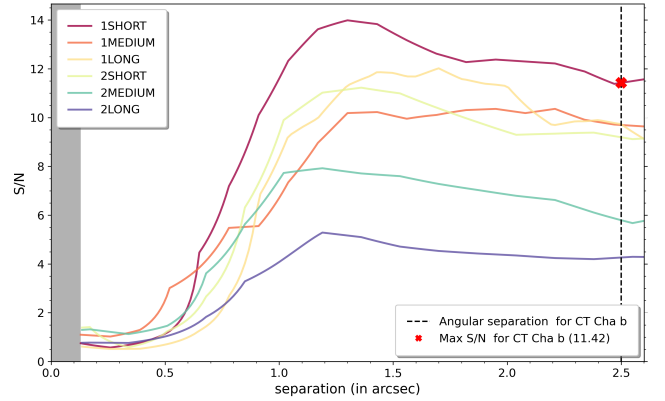
discrepancy between the observed spectrum and the template ( $\cos \theta_p$ ). We estimate the quantity in Eq. (17) for different S/Ns per spectral channel and plot it in Fig. 18. It illustrates that the greater the S/N per spectral channel, the closer the estimated correlation will be to the mismatch  $\cos \theta_p$  (or to  $\cos \theta_{\text{lim}} \times \cos \theta_p$  if there are systematic modulations). It is also observed that the greater the planet's temperature, the greater the S/N per spectral channel required to retrieve the mismatch. This is simply because the higher the temperature, the lower the spectral content and the less it will be distinguishable from the noise.

If self-subtraction is now considered, it is written:

$$\cos \theta_{\text{est}} = \cos \theta_n \times \cos \theta_{\text{lim}} \times \cos \theta_p - \frac{\beta}{\|\hat{d}(\lambda)\|} \quad (18)$$

$$\rightarrow \cos \theta_p = \frac{\cos \theta_{\text{est}} / \cos \theta_n + \beta / \alpha}{\cos \theta_{\text{lim}}},$$

where  $\beta/\alpha$  is the fraction of the signal that is self-subtracted. Hence, a straightforward and linear relationship exists between the estimated correlation and the mismatch. The best that can be done now is to assume that, even if the spectrum model differs from the observed one, the quantities  $\cos \theta_n$ ,  $\cos \theta_{\text{lim}}$ , and  $\beta/\alpha$  are accurately estimated with FastCurves (knowing that the noise level is well estimated). This way, the estimated correlation of around 0.29 actually implies a mismatch of around



**Fig. 19.** S/N curves estimated with FastCurves for the CT Cha b observation with MIRI/MRS (with  $t_{\text{exp}} = 56$  min,  $K_s = 8.7$ ,  $K_p = 14.9$ , and with the BT-Settl spectrum at 2600 K).

0.82 with the template used (with  $\cos \theta_{\text{est}} \approx 0.29$ ,  $\cos \theta_n \approx 0.54$ ,  $\cos \theta_{\text{lim}} \approx 0.94$ , and  $\beta/\alpha \approx 0.23$  at the planet separation). It is noteworthy that if the template was exactly the observed spectrum ( $\cos \theta_p = 1$ ), a correlation of about 0.38 would have been measured. The model used therefore seems to be much more similar to the observed spectrum than it initially appeared. Lastly, the S/N curves for the case of CT Cha b were calculated by introducing a mismatch of 0.82 (see Fig. 19), revealing that FastCurves estimates a S/N of around 11.4 on 1SHORT, which closely aligns with the observed S/N (10.5, see upper right panel of Fig. 6). Similar observations hold true for the other bands of channel 1. This underscores the capability of FastCurves to reliably estimate the signal as well.

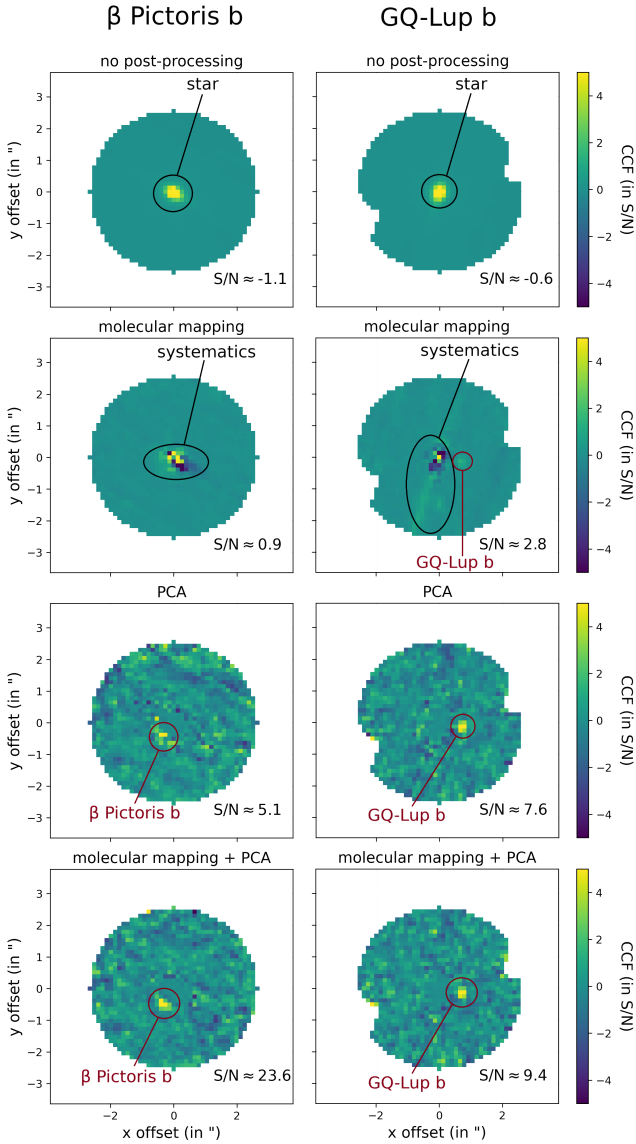
## 7. Enhancing detection limits with PCA

Given the presence of systematics, it is worth attempting to mitigate their impact by applying a PCA to the spectral dimension of the data. Although PCA is typically used with temporal diversity, here it leverages the spectral diversity of spaxels within the data cube to subtract principal modes associated with spatially varying systematic spectral modulations. This application of PCA to the spectral dimension has been shown to be effective for exoplanet detection, as has been demonstrated by Ruffio et al. (2024), Parker et al. (2024), and Landman et al. (2024). We applied PCA to the filtered cube,  $S_{\text{res}}$  (see Eq. (B.6)), and then performed cross-correlation, making both  $\beta$  Pictoris b<sup>10</sup> and GQ-Lup b<sup>11</sup> detectable (see Fig. 20). Detailed analyses of these detections are provided in Worthen et al. (2024) and Cugno et al. (2024), respectively.

Focusing on  $\beta$  Pictoris b, we calculated the S/N for various post-processing parameters, such as varying cutoff resolutions and the number of subtracted principal components (see Figs. 21 and 22). Additionally, we incorporated PCA into FastCurves predictions by applying it to the simulated noiseless cubes, which we used to assess the level of systematics and estimate the corresponding signal drop. In the middle panel of Fig. 21, we show the contrasts estimated from the data that were obtained using the  $\alpha_0$  calculated by FastCurves, allowing for a comparison of the different noise levels in the first instance. It is observed that

<sup>10</sup> PI: C. Chen, <https://www.stsci.edu/jwst/phase2-public/1294.pdf>

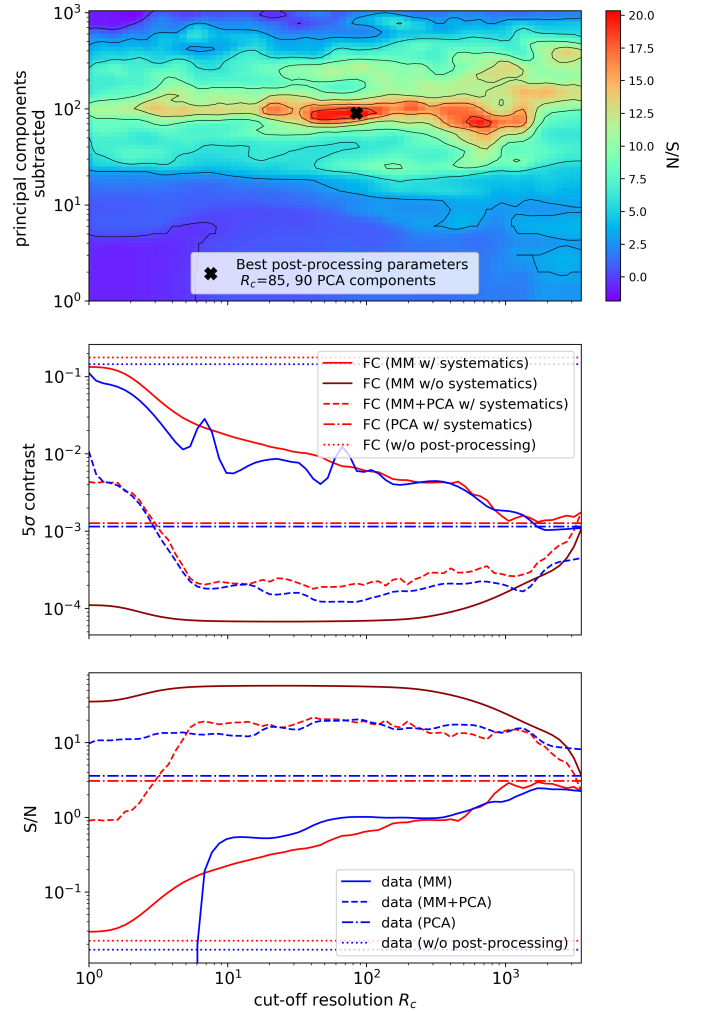
<sup>11</sup> PI: A. Banzatti, <https://www.stsci.edu/jwst/phase2-public/1640.pdf>



**Fig. 20.** CCF S/N maps on 1SHORT for  $\beta$  Pictoris b data (with  $t_{exp} = 50$  min,  $K_* = 3.5$ ,  $K_p \approx 12.7$  and  $T_{eff} \approx 1600$  K) and GQ-Lup b (with  $t_{exp} = 30$  min,  $K_* = 7.1$ ,  $K_p \approx 13.5$  and  $T_{eff} \approx 2700$  K) using various post-processing techniques (the S/N color scales are cropped to  $\pm 5$ ). Molecular mapping was applied with a cutoff resolution of 100, and PCA was used to subtract 100 principal components.

the noise levels estimated by FastCurves still align well with those in the data for the same post-processing methods. Similarly, the estimated and measured signal levels are consistent (bottom panel of Fig. 21). In this case, PCA alone outperforms molecular mapping alone, but the best detection performance is achieved by combining PCA with molecular mapping. This combined approach closely approaches the performance set by fundamental noise limits, though it does not completely reach that level due to the signal self-subtraction induced by PCA. The positive impact of PCA is further highlighted in Fig. 22, which shows its critical role in detecting close companions in systematics-dominated regimes.

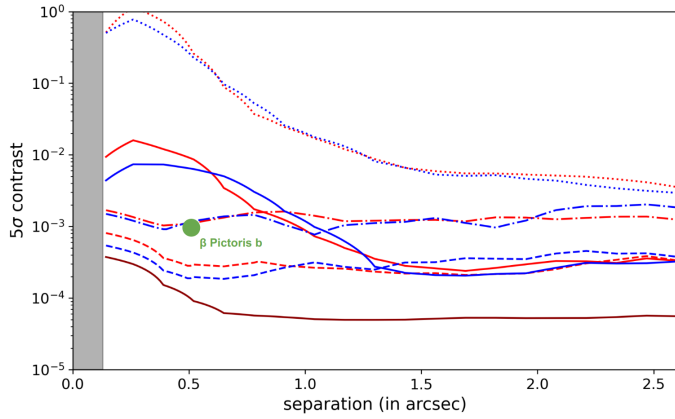
The optimal choice of the number of subtracted principal components is a balance between reducing systematic noise and minimizing signal loss (see Fig. 23). This optimal number primarily depends on the planet’s separation and the contrast



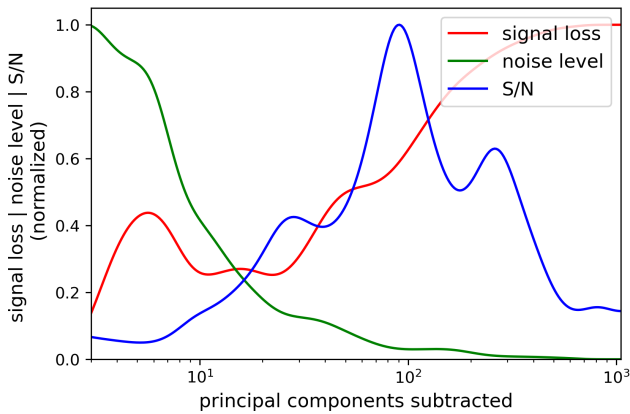
**Fig. 21.** Top: S/N fluctuation map as a function of post-processing parameters on 1SHORT for  $\beta$  Pictoris b on-sky data. Middle: contrast curves calculated analytically with FastCurves (in red) and empirically with  $\beta$  Pictoris b on-sky data (in blue) on 1SHORT for various post-processing techniques. Bottom: S/N curves calculated similarly. The plots are given at the planet’s separation ( $\rho \approx 0.5''$ ) and for 90 principal components subtracted when the PCA is applied. For FastCurves estimations, as in Fig. 9, systematics ( $M_{sim}$ ) where fringes and straylight were incorrectly subtracted by the pipeline had to be considered (see Sect. 5).

between the planet and the star, which determines how much of the planet’s spectrum enters the principal components, and thus the degree of self-subtraction of the signal. For  $\beta$  Pictoris b, this optimal number is around 100 components, resulting in a signal reduction of about 60% but a systematic noise reduction of approximately 95%.

Finally, examining the modulation patterns of the first principal components estimated by the PCA on the filtered cube ( $S_{res}$ ), along with the frequencies they influence and their projection by correlation on  $S_{res}$ , provides valuable insights (see Fig. 24). It is observed that the first two modes are predominantly affected by fringing (evident from their frequency peaks in the PSDs), while subsequent modes begin to reflect modulations caused by aliasing artifacts (with peaks at lower frequencies around  $R \approx 100$ ), consistent with previous observations. Furthermore, the correlation projection of these modes on  $S_{res}$  reveals significant correlation intensity (up to approximately 0.8) at the stellar



**Fig. 22.** Same as the middle panel of Fig. 21 (using the same color scheme) but presented as a function of separation instead of cutoff resolution. The plots correspond to the optimal post-processing parameters for  $\beta$  Pictoris b on-sky data in 1SHORT, with a cutoff resolution of 85 and the subtraction of 90 principal components.



**Fig. 23.** Trade-off on the optimal number of principal components subtracted for  $\beta$  Pictoris b on 1SHORT when combining molecular mapping (with  $R_c = 100$ ) and PCA.

position, along with diagonal and parallel structures typical of straylight effects. Essentially, straylight would not be as limiting without the presence of high-frequency modulations induced by fringing and aliasing. However, straylight scatters stellar flux away from the star, resulting in straight, parallel structures that enhance the effects of fringe- and aliasing-induced modulations.

It is important to note that PCA appears effective in both cases ( $\beta$  Pictoris b and GQ-Lup b), as detection is otherwise limited by systematics. However, when systematics are not dominant, such as in short exposure times, low stellar flux, or high separation, PCA will not be relevant.

## 8. Conclusions

In this study, we compare the effective detection limits for close companions, derived from real-life IFS data processed with molecular mapping, to the limits obtained from the semi-analytical FastCurves algorithm. Molecular mapping is theoretically capable of eliminating the detection limits imposed by speckle noise, potentially yielding photon noise-limited data unless other noise sources, related to nonideal instrument effects, become dominant. We selected JWST/MIRI/MRS IFU data for

this comparison due to their sufficiently high spectral resolution, the absence of AO residuals, and the ability to compare these results with predictions from MIRISim, the MIRI/MRS simulation tool.

### 8.1. Understanding systematic effects and their impact

We have accurately characterized the fundamental noises of MIRI/MRS with the analytical tool, as the FastCurves predictions align with the data ERR extension. Noise propagation through the CCF is also well captured, with agreement between analytical predictions from FastCurves and empirical spatial variance on CCF maps. Since the noise in the CCF cannot be explained by fundamental noise sources alone, it becomes essential to incorporate a systematic noise component proportional to the stellar flux. This component is well characterized through tests on MIRISim and the pipeline, aligning closely with real data and attributed to nonideal instrument effects and signal projection and retrieval on the detector, setting an ultimate contrast limit that cannot be overcome solely with molecular mapping.

We have identified and explored the detection limits of MIRI/MRS and their underlying sources. Various origins of systematic noise were investigated, excluding factors such as bad outlier subtraction, different 3D reconstruction algorithms, or wavelength calibration errors. The study highlights the significant impact of straylight, fringing, and resampling noise on spatial noise statistics. Systematic effects during signal extraction can also decrease the signal correlation from observed planets. Adjusting the cutoff resolution enhances performance under systematic conditions, and exposure durations must be meticulously selected to optimize observation efficiency given these effects.

### 8.2. Implications of noise propagation and template assumptions

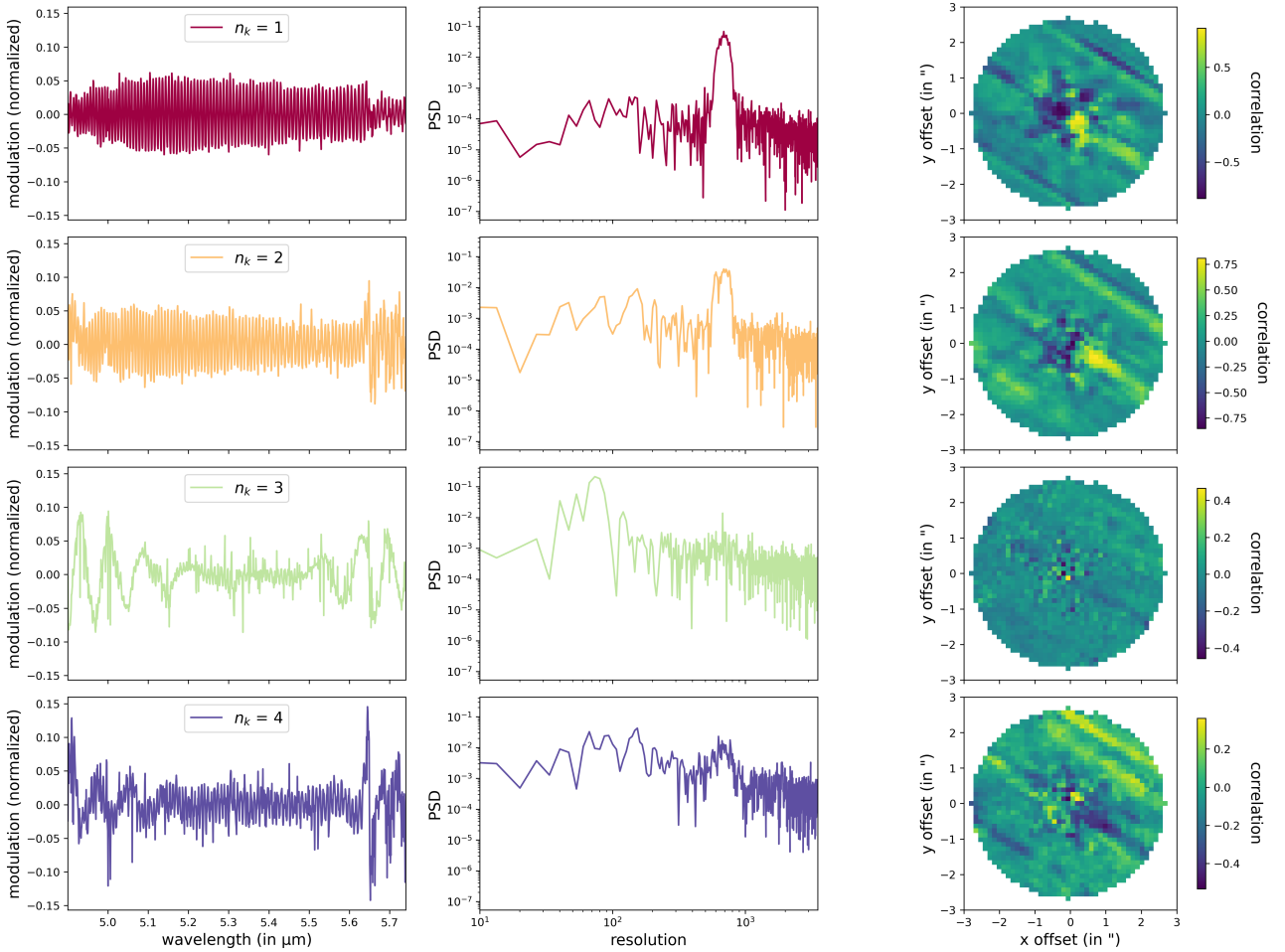
We have discussed the implications of noise propagation and template assumptions in interpreting planetary spectrum correlations with assumed models, ensuring accurate astronomical interpretations from correlation intensity. In particular, systematic noise not only impacts detection sensitivity but can also introduce biases in inferred planetary properties if not properly accounted for. Understanding the interplay between noise characteristics, template accuracy, and signal retrieval methods is essential to derive robust planetary spectra cross-correlation techniques.

### 8.3. PCA as a key tool for systematic noise mitigation

A major result of this study is the demonstrated effectiveness of PCA in mitigating systematic noise. By leveraging spectral diversity, PCA removes dominant systematic components and enhances detection sensitivity. In the case of  $\beta$  Pictoris b, PCA increases the S/N by a factor of  $\sim 25$ , making it a crucial technique for extracting planetary signals. However, PCA also induces signal loss, requiring careful tuning of the number of subtracted components to optimize detection performance.

### 8.4. Recommendations for future observations and data processing

Building on these findings, we propose an optimized observing and data processing strategy for molecular mapping with FastCurves:



**Fig. 24.** First four modes of the PCA performed on the filtered cube  $S_{\text{res}}$  of the  $\beta$  Pictoris b data on ISHORT (left), their PSD (middle), and their correlation projection onto  $S_{\text{res}}$  (right).

- Performance assessment: we recommend using FastCurves in advance to evaluate whether the chosen instrument, considering its known systematic effects, can achieve the scientific objectives within a reasonable exposure time;
- Exposure time optimization: we emphasize the need to carefully plan exposure time, as beyond a certain threshold the S/N no longer increases linearly, leading to diminishing returns and reducing the instrument’s efficiency. This aspect is crucial when optimizing telescope time allocation;
- Post-processing parameter optimization: after data acquisition, the optimal cutoff resolution should be selected to enhance planetary signal detection while efficiently suppressing stellar speckle noise and systematic modulations. Applying PCA to the filtered data cube significantly improves the S/N, but the number of principal components subtracted must be carefully tuned to maximize systematic noise suppression while minimizing signal loss, following the same principle as the cutoff resolution selection;
- Systematic noise mitigation in instrument design: future instrument development must incorporate constraints on systematic noise to ensure the continued effectiveness of cross-correlation techniques for exoplanet detection and characterization. Additionally, we stress the importance of assessing potential systematics and anticipating precise calibrations to minimize their impact on overall performance.

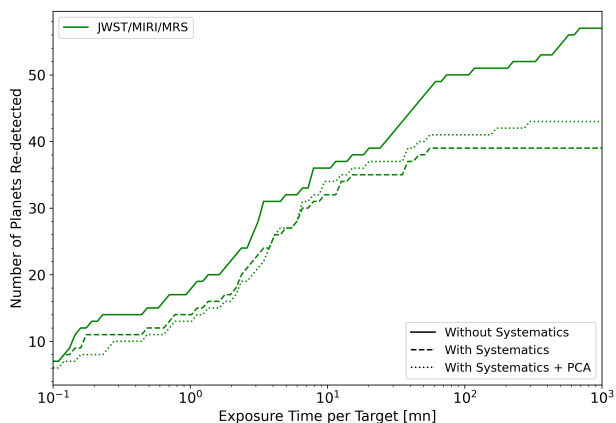
### 8.5. Prospects for future instruments

While this study focuses on MIRI/MRS, the methodology and insights gained are broadly applicable to other instruments, including ground-based IFSs such as ELT/HARMONI, ELT/METIS, ELT/ANDES, or ELT/PCS, as well as future space-based observatories. The refined noise models and systematic noise mitigation techniques presented here will be essential for defining noise budgets, optimizing observational strategies, and improving exoplanet characterization capabilities across a range of instruments.

### 8.6. Final remarks

This work demonstrates the feasibility of predicting instrument performance when combining spectroscopy with high-contrast imaging while incorporating systematic effects. By verifying that noise and signal levels in real data match model predictions, we have developed an extension of FastCurves<sup>12</sup> to assess the detection yield of an instrument across an exoplanet catalog. Applying this tool to an archive of known exoplanets reveals how systematic effects impact instrument performance and potentially limit detections (see Fig. 25). This highlights the crucial

<sup>12</sup> <https://github.com/StevMartos/FastYield>



**Fig. 25.** MIRI/MRS known exoplanets detection yields with molecular mapping. Yields were estimated by applying the FastCurves S/N calculation (i.e., Fig. 19) to each of the 5600 planets in the Planetary Systems and Planetary Systems Composite Parameters Table from the NASA exoplanet archive (Akeson et al. 2013). A planet is considered detected when the S/N on one of the MIRI/MRS bands is greater than 5 for the exposure time considered. It can be observed that PCA would allow MIRI/MRS to approach the performance levels it could achieve without systematics, though not fully reach them, due to the signal loss caused by the subtraction of the principal components.

role of systematics in defining detection limits and shaping the future of exoplanet characterization.

*Acknowledgements.* The authors extend their gratitude to Benjamin Charnay for generously sharing the Exo-REM spectra grid with a resolution of  $R = 200000$ , which proved more suitable for the MIRI/MRS case compared to the publicly available grid ( $R = 20000$  at  $1 \mu\text{m}$ ). Additionally, they express appreciation to Mathilde Mâlin for providing the simulated data cubes used in her article, and to David Law for engaging discussions regarding resampling noise and the projection of noise statistics through the pipeline's 3D cube reconstruction algorithm. This project receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 866001 – EXACT).

## References

- Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, *PASP*, **125**, 989
- Allard, F., Homeier, D., & Freytag, B. 2012, *Philos. Trans. R. Soc. Lond. A*, **370**, 2765
- Argyriou, I., Wells, M., Glasse, A., et al. 2020, *A&A*, **641**, A150
- Argyriou, I., Glasse, A., Law, D. R., et al. 2023, *A&A*, **675**, A111
- Batalha, N. E., Marley, M. S., Lewis, N. K., & Fortney, J. J. 2019, *ApJ*, **878**, 70
- Beuzit, J.-L., Vigan, A., Mouillet, D., et al. 2019, *A&A*, **631**, A155
- Bidot, A., Mouillet, D., & Carloti, A. 2024, *A&A*, **682**, A10
- Boccaletti, A., Mâlin, M., Baudoz, P., et al. 2024, *A&A*, **686**, A33
- Bonnefoy, M., Chauvin, G., Lagrange, A.-M., et al. 2014, *A&A*, **562**, A127
- Charnay, B., Bézard, B., Baudino, J.-L., et al. 2018, *ApJ*, **854**, 172
- Cugno, G., Patapis, P., Banzatti, A., et al. 2024, *ApJ*, **966**, L21
- Delorme, J.-R., Jovanovic, N., Echeverri, D., et al. 2021, *JATIS*, **7**, 035006
- Deming, D., Fu, G., Bouwman, J., et al. 2024, *PASP*, **136**, 084402
- Fruchter, A. S., & Hook, R. N. 2002, *PASP*, **114**, 144
- Garvin, E. O., Bonse, M. J., Hayoz, J., et al. 2024, *A&A*, **689**, A143
- Gasman, D., Argyriou, I., Sloan, G. C., et al. 2023, *A&A*, **673**, A102
- Gasman, D., Argyriou, I., Morrison, J. E., et al. 2024, *A&A*, **688**, A226
- Henning, T., Kamp, I., Samland, M., et al. 2024, *PASP*, **136**, 054302
- Hoeijmakers, H. J., Schwarz, H., Snellen, I. A. G., et al. 2018, *A&A*, **617**, A144
- Jorquera, S., Bonnefoy, M., Pérez, L. M., et al. 2024, *ApJ*, **976**, 42
- Jovanovic, N., Martinache, F., Guyon, O., et al. 2015, *PASP*, **127**, 890
- Klaassen, P. D., Geers, V. C., Beard, S. M., et al. 2020, *MNRAS*, **500**, 2813
- Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, *Science*, **339**, 1398
- Kravchenko, K., Dallilar, Y., Absil, O., et al. 2022, in *Ground-based and Airborne Instrumentation for Astronomy IX*, 9908, eds. C. J. Evans, J. J. Bryant, & K. Motohara, (SPIE), 211
- Labiano-Ortega, A., Dicken, D., Vandenbussche, B., et al. 2016, in *Observatory Operations: Strategies, Processes, and Systems VI*, eds. A. B. Peck, C. R. Benn, & R. L. Seaman (SPIE)
- Landman, R., Snellen, I. A. G., Keller, C. U., et al. 2023, *A&A*, **675**, A157
- Landman, R., Stolker, T., Snellen, I., et al. 2024, *A&A*, **682**, A48
- Law, D. R., E. Morrison, J., Argyriou, I., et al. 2023, *AJ*, **166**, 45
- Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011, *Nature*, **470**, 53
- Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, *PNAS*, **111**, 12661
- Mâlin, M., Boccaletti, A., Charnay, B., Kiefer, F., & Bézard, B. 2023, *A&A*, **671**, A109
- Marley, M. S., Saumon, D., Visscher, C., et al. 2021, *ApJ*, **920**, 85
- Miles, B. E., Biller, B. A., Patapis, P., et al. 2023, *ApJL*, **946**, L6
- Mollière, P., Molyarova, T., Bitsch, B., et al. 2022, *ApJ*, **934**, 74
- Morley, C. V., Fortney, J. J., Marley, M. S., et al. 2012, *ApJ*, **756**, 172
- Mâlin, M., Boccaletti, A., Perrot, C., et al. 2024, *A&A*, **690**, A316
- Oberg, K. I., & Bergin, E. A. 2016, *ApJ*, **831**, L19
- Parker, L. T., Birkby, J. L., Landman, R., et al. 2024, *MNRAS*, **531**, 2356
- Patapis, P., Nasedkin, E., Cugno, G., et al. 2022, *A&A*, **658**, A72
- Patience, J., King, R. R., De Rosa, R. J., et al. 2012, *A&A*, **540**, A85
- Perrin, M., Sivaramakrishnan, A., Makidon, R., Oppenheimer, B., & Graham, J. 2003, *ApJ*, **596**
- Petit dit de la Roche, D. J. M., Hoeijmakers, H. J., & Snellen, I. A. G. 2018, *A&A*, **616**, A146
- Petrus, S., Bonnefoy, M., Chauvin, G., et al. 2021, *A&A*, **648**, A59
- Pontoppidan, K. M., Salyk, C., Banzatti, A., et al. 2024, *ApJ*, **963**, 158
- Rieke, G. H., Ressler, M. E., Morrison, J. E., et al. 2015, *PASP*, **127**, 665
- Rigby, J. R., Lightsey, P. A., García Marín, M., et al. 2023, *PASP*, **135**, 048002
- Ruffio, J.-B., Macintosh, B., Konopacky, Q. M., et al. 2019, *AJ*, **158**, 200
- Ruffio, J.-B., Perrin, M. D., Hoch, K. K. W., et al. 2024, *AJ*, **168**, 73
- Schmidt, T. O. B., Neuhäuser, R., Seifahrt, A., et al. 2008, *A&A*, **491**, 311
- Smith, J. D. T., Armus, L., Dale, D. A., et al. 2007, *PASP*, **119**, 1133
- Snellen, I., de Kok, R., Birkby, J. L., et al. 2015, *A&A*, **576**, A59
- Turbet, M. 2018, Ph.D. Thesis, Sorbonne Université, France
- Vigan, A., Morsy, M. E., Lopez, M., et al. 2024, *A&A*, **682**, A16
- Wang, J., Mawet, D., Hu, R., et al. 2018, *JATIS*, **4**, 035001
- Wells, M., Pel, J.-W., Glasse, A., et al. 2015, *PASP*, **127**, 646
- Worthen, K., Chen, C. H., Law, D. R., et al. 2024, *ApJ*, **964**, 168
- Wu, Y.-L., Close, L. M., Males, J. R., et al. 2015, *ApJ*, **801**, 4

## Appendix A: Used filter

The high-pass filtering operation was calculated according to

$$[S(\lambda)]_{\text{HF}} = S(\lambda) - [S(\lambda)]_{\text{LF}}. \quad (\text{A.1})$$

Here, `scipy`'s Gaussian filtering (`gaussian_filter`<sup>13</sup>) was used. The low-pass filtering operation is a convolution of the considered spectrum and a Gaussian kernel of an arbitrarily defined width,  $\sigma_c$ . This width determines the cutoff frequency,  $f_c$ , and the cutoff resolution,  $R_c$ :

$$[S(\lambda)]_{\text{LF}} = S(\lambda) * G_{\sigma_c}(\lambda), \text{ with } G_{\sigma_c}(\lambda) = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{-\frac{\lambda^2}{2\sigma_c^2}}. \quad (\text{A.2})$$

If the focus is on the PSD of the spectrum, it can be written as

$$PSD(f) = |\text{TF}\{S(\lambda)\}(f) \times \text{TF}\{G_{\sigma_c}(\lambda)\}(f)|^2. \quad (\text{A.3})$$

The cutoff frequency,  $f_c$ , is defined as

$$\text{TF}\{G_{\sigma_c}(\lambda)\}(f_c) = \frac{1}{2}. \quad (\text{A.4})$$

The Fourier transform of the Gaussian kernel is given by

$$\text{TF}\{G_{\sigma_c}(\lambda)\}(f) = e^{-2\pi^2 f^2 \sigma_c^2}. \quad (\text{A.5})$$

So,

$$e^{-2\pi^2 f_c^2 \sigma_c^2} = \frac{1}{2} \rightarrow f_c = \frac{1}{\pi\sigma_c} \sqrt{\frac{\ln(2)}{2}}. \quad (\text{A.6})$$

The cutoff frequency,  $f_c$ , is then converted into the cutoff resolution,  $R_c$ , by

$$R_c = R_{\text{sampling}} \times f_c = 2R_{\text{instru}} \times f_c, \quad (\text{A.7})$$

where  $R_{\text{sampling}}$  is the sampling frequency and  $R_{\text{instru}}$  is the instrument resolution (assuming Nyquist sampling), with

$$R_{\text{instru}} = \frac{\lambda}{2\Delta\lambda} \text{ and } R_{\text{sampling}} = \frac{\lambda}{\Delta\lambda}. \quad (\text{A.8})$$

Finally,

$$\sigma_c = \frac{2R_{\text{instru}}}{\pi R_c} \sqrt{\frac{\ln(2)}{2}}. \quad (\text{A.9})$$

## Appendix B: Molecular mapping formalism (with systematics)

To isolate the planetary spectrum from the stellar spectrum, we considered the `TexTris` pre-processing routine (Petrus et al. 2021). First, the spectrum of the star was empirically estimated from the data:

$$\hat{\gamma}(\lambda)\hat{S}_*(\lambda) = \sum_{i,j} S(\lambda, x_i, y_j). \quad (\text{B.1})$$

Then, the stellar modulation function was estimated with

$$\hat{M}(\lambda, x, y) = \left[ \frac{S(\lambda, x, y)}{\hat{\gamma}(\lambda)\hat{S}_*(\lambda)} \right]_{\text{LF}}. \quad (\text{B.2})$$

Next, the stellar component was subtracted from the above-estimated quantities:

$$S_{\text{res}}(\lambda, x, y) = S(\lambda, x, y) - \hat{M}(\lambda, x, y)\hat{\gamma}(\lambda)\hat{S}_*(\lambda). \quad (\text{B.3})$$

It was assumed that  $\hat{\gamma}\hat{S}_* \approx \gamma S_*$  and that  $[S \times S']_{\text{LF}} \approx [S]_{\text{LF}} \times [S']_{\text{LF}}$ . Thus, the estimated stellar modulation function can be written as

$$\begin{aligned} \hat{M}(\lambda, x, y) &= \left[ \frac{S(\lambda, x, y)}{\hat{\gamma}(\lambda)\hat{S}_*(\lambda)} \right]_{\text{LF}} \\ &= [M(\lambda, x, y)]_{\text{LF}} + \left[ \frac{M_p(\lambda, x, y)S_p(\lambda)}{S_*(\lambda)} \right]_{\text{LF}} \\ &\quad + \left[ \frac{n(\lambda, x, y)}{\gamma(\lambda)S_*(\lambda)} \right]_{\text{LF}}. \end{aligned} \quad (\text{B.4})$$

So:

$$\begin{aligned} S_{\text{res}}(\lambda, x, y) &= S(\lambda, x, y) - \hat{M}(\lambda, x, y)\hat{\gamma}(\lambda)\hat{S}_*(\lambda) \\ &= [M(\lambda, x, y)]_{\text{HF}}\gamma(\lambda)S_*(\lambda) + \gamma(\lambda)M_p(\lambda, x, y)S_p(\lambda) \\ &\quad + n(\lambda, x, y) - \left[ \frac{M_p(\lambda, x, y)S_p(\lambda)}{S_*(\lambda)} \right]_{\text{LF}}\gamma(\lambda)S_*(\lambda) \\ &\quad - \left[ \frac{n(\lambda, x, y)}{\gamma(\lambda)S_*(\lambda)} \right]_{\text{LF}}\gamma(\lambda)S_*(\lambda). \end{aligned} \quad (\text{B.5})$$

Bidot et al. (2024) have shown that the last low-frequency filtered noise term  $[n/\gamma S_*]_{\text{LF}}\gamma S_*$  is negligible compared with the noise term  $n$ . Then:

$$\begin{aligned} S_{\text{res}}(\lambda, x, y) &= \gamma(\lambda)[M_p(\lambda, x, y)S_p(\lambda)]_{\text{HF}} \\ &\quad - \left[ \frac{M_p(\lambda, x, y)S_p(\lambda)}{S_*(\lambda)} \right]_{\text{LF}}\gamma(\lambda)[S_*(\lambda)]_{\text{HF}} \\ &\quad + n(\lambda, x, y) + n_{\text{sys}}(\lambda, x, y), \end{aligned} \quad (\text{B.6})$$

with

$$n_{\text{sys}}(\lambda, x, y) = [M(\lambda, x, y)]_{\text{HF}}\gamma(\lambda)S_*(\lambda). \quad (\text{B.7})$$

This way, a cube  $S_{\text{res}}$  is obtained in which the stellar contribution modulated by  $M$  was estimated and subtracted.

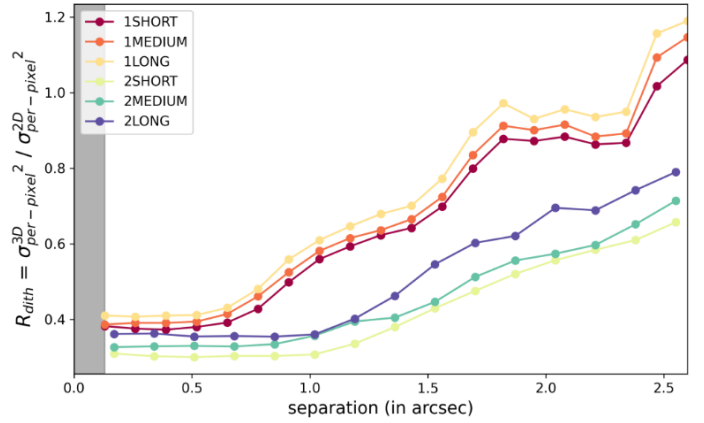
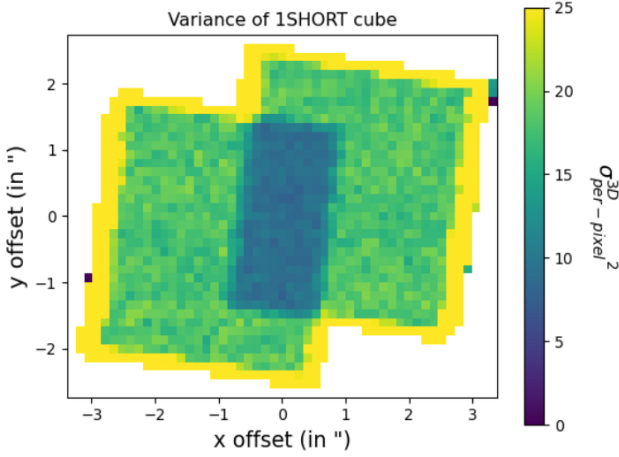
If  $\mathbb{S}_p$  is the planetary spectrum model considered, the template used is this same spectrum but normalized, high-pass filtered, and shifted with a radial velocity,  $rv$ ; that is,

$$\hat{t}(\lambda, rv) = \frac{\gamma(\lambda)[\mathbb{S}_p(\lambda, rv)]_{\text{HF}}}{\|\gamma(\lambda)[\mathbb{S}_p(\lambda, rv)]_{\text{HF}}\|}. \quad (\text{B.8})$$

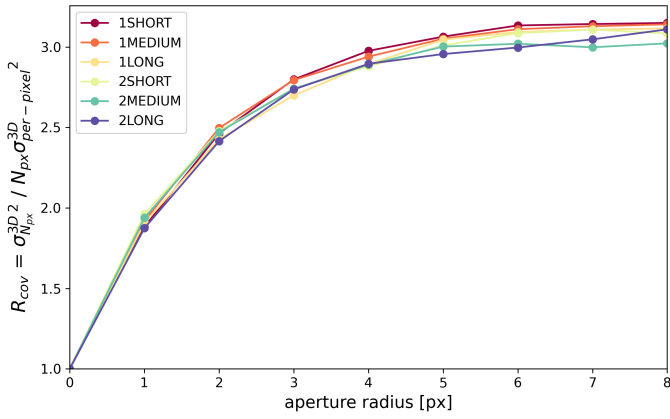
To simplify the notation, the radial velocity dependence,  $rv$ , is omitted, but it should be kept in mind that the observed stellar and planetary spectra depend on their respective radial velocities. The 2D CCF is simply given by the scalar product between the residual signal and the template, such as

$$\begin{aligned} \text{CCF}(x, y) &= \langle S_{\text{res}}(\lambda, x, y), \hat{t}(\lambda) \rangle \\ &= \|\gamma(\lambda)[M_p(\lambda, x, y)S_p(\lambda)]_{\text{HF}}\| \cos \theta_p \\ &\quad - \left\| \left[ \frac{M_p(\lambda, x, y)S_p(\lambda)}{S_*(\lambda)} \right]_{\text{LF}} \gamma(\lambda)[S_*(\lambda)]_{\text{HF}} \right\| \cos \theta_* \\ &\quad + \langle n(\lambda, x, y) + n_{\text{sys}}(\lambda, x, y), \hat{t}(\lambda) \rangle \end{aligned} \quad (\text{B.9})$$

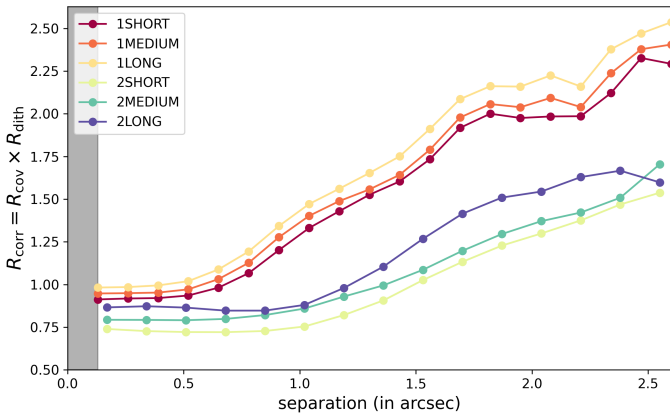
<sup>13</sup> [https://github.com/scipy/scipy/blob/v1.12.0/scipy/ndimage/\\_filters.py](https://github.com/scipy/scipy/blob/v1.12.0/scipy/ndimage/_filters.py)



**Fig. C.1.** Left: variance (in the spectral dimension) of a cube on the 1SHORT-band, while injecting an arbitrarily defined variance. Right: ratios between the per-pixel variances in the 3D cubes and the per-pixel variances in the 2D detector images used to reconstruct the cubes.



**Fig. C.2.** Ratio between the measured variance in a given aperture radius and the variance that would be measured without covariance between spaxels.



**Fig. C.3.** Ratios between the per-FWHM variances in the 3D cubes and the per-FWHM variances in the 2D detector images used to reconstruct the cubes.

### Appendix C: Corrective factor $R_{\text{corr}}$ in dithered data

It is fairly straightforward to estimate the noise that 2D detector images of MIRI/MRS should theoretically have, given the (Gaussian) photon and detector noises. However, the drizzle

method of reconstruction to a 3D cube (Law et al. 2023) modifies the statistics of the noise.

We let  $\sigma_{\text{per-pixel}}^{2D}$  be the variance per-pixel of the 2D detector image considered (this is the variance calculated with FastCurves for photon and detector noises) and  $\sigma_{\text{per-pixel}}^{3D}$  be the variance actually found in the reconstructed cube. The ratio between these two quantities is called  $R_{\text{dith}}$ . To measure this quantity empirically, the values of four 2D detector images corresponding to four different dithering points are replaced by values for which the variance is arbitrarily defined. Then, the data cubes are reconstructed with the pipeline and the effective variances measured. But, as is shown on the left panel of the Fig. C.1, the dithering is not spatially homogeneous. In particular, the center is a region where all four integrations of the four-point dithering are being stacked, while at the edges only two out of four will be stacked, giving a variance two times greater than at the center. This implies the need to calculate  $R_{\text{dith}}$  for each separation (see right panel of Fig. C.1). If there is no obligation to place the star in the center and the planet's position is known, it is then advisable to place the planet in the central region of the FoV, where all the dithering positions will be stacked.

In the method considered, the planet's signal is integrated over a box of  $A_{\text{fwhm}}$  spaxels and the dithering implies spatial covariance between adjacent spaxels, so the variance in this box will not be equal to  $A_{\text{fwhm}} \times \sigma_{\text{per-pixel}}^{3D}$ . We defined

$$R_{\text{cov}} = \frac{\sigma_{N_{\text{px}}}^{3D}}{N_{\text{px}} \sigma_{\text{per-pixel}}^{3D}}, \quad (\text{C.1})$$

where  $\sigma_{N_{\text{px}}}^{3D}$  is the variance measured when summing the signal over an aperture of  $N_{\text{px}}$  pixels. This calculation is similar to the one performed in Law et al. (2023). Figure C.2 shows that this ratio increases asymptotically with the aperture radius due to the covariance. Nevertheless, the interest is on an aperture of  $A_{\text{fwhm}}$  spaxels, so:  $R_{\text{cov}} = \sigma_{A_{\text{fwhm}}}^{3D} / A_{\text{fwhm}} \sigma_{\text{per-pixel}}^{3D}$ , where  $\sigma_{A_{\text{fwhm}}}^{3D}$  is the effective variance when summing the signal over an aperture of  $A_{\text{fwhm}}$  spaxels (this is the quantity that must be expressed in terms of  $\sigma_{\text{per-pixel}}^{2D}$ ). Finally:

$$\sigma_{A_{\text{fwhm}}}^{3D} = A_{\text{fwhm}} R_{\text{corr}} \sigma_{\text{per-pixel}}^{2D}, \quad (\text{C.2})$$

where  $R_{\text{corr}} = R_{\text{cov}} \times R_{\text{dith}}$  is the corrective ratio introduced in the core of the paper. This quantity is plotted for each band in Fig. C.3.

## Appendix D: Structure and spatial variations in the modulations through PSD

To assess the impact of various effects, the previous method used to obtain the simulated stellar modulation function  $M_{\text{sim}}$  is repeated, but with a choice to include or exclude specific corrections. In particular, fringes can be corrected or not, since MIRISim simulates the fringes with the same extended-source fringe pattern as the 2D flat used in the first pipeline correction step. This is why only the first fringe subtraction step is necessary to perfectly eliminate all fringes from the data cubes simulated by MIRISim. In addition, bad pixels, hot pixels, and cosmic rays (outliers) are absent from the noiseless 2D detector images simulated with MIRISim. Therefore, applying the outliers subtraction step may yield to a modulation function assuming bad outliers subtraction. Similarly, as MIRISim does not simulate straylight, applying the straylight subtraction step will result in a modulation function with bad straylight subtraction. Finally, an Exponential Modified Shepard Method (EMSM) can be used to reconstruct the cube, also detailed in Law et al. (2023). This enables a comparison between the two cube reconstruction methods, drizzle and EMSM, to identify any discrepancies. Although the focus here is on the 1SHORT band, similar outcomes can be derived for other bands as well.

To get an idea of the impact of these different effects and corrections on the modulation structure, the PSD of  $[M_{\text{sim}}]_{\text{HF}}$  is computed for each spaxel and averaged for each separation (see left panel of Fig. D.1). Even with all the corrections applied (top left), residual modulations remain after high-pass filtering, peaking at  $R \approx 100$ . This peak primarily arises from residual aliasing artifacts, and shows why it makes sense to filter at a higher cut-off resolution to mitigate systematics (Sect. 4.3). In addition, a different reconstruction method (middle left) or a bad outliers subtraction (middle right) does not seem to imply any differences in modulations. If fringes are left uncorrected (top right), an additional peak emerges in the PSD at  $R \approx 800$ , corresponding to the frequency of simulated fringes. Conversely, bad straylight subtraction (bottom right) leads to a drastic alteration in the PSD, affecting nearly all resolutions and compromising fringe suppression, even with the correct flat applied. Furthermore, a significant difference is found between the PSDs of stellar modulation functions estimated from simulations (top left) and data (bottom left) when all corrections are applied. While fringes persist in the data, as expected with only the first fringe correction step applied, they occur at a different resolution ( $R \approx 650$ ) compared to simulated fringes. There are also additional modulations in the data at high-resolution, stemming from bad straylight subtraction (directly visible in the data), but to a lesser extent than simulated (bottom right). Lastly, it can be seen from Fig. D.1 that the spatial standard deviations of the PSDs are found to be equivalent to their means.

## Appendix E: Simulated systematic effects on noise

### E.1. Photometric calibration errors

The impact of photometric calibration errors, akin to resampling noise, on contrast is quantified using the following approach. A low-pass filter with a strong cutoff resolution ( $R_c = 10$ ) is applied on  $M_{\text{sim}}$  (with all corrections) to retrieve a stellar

modulation function without any systematic high-frequency modulations (only speckles). On the latter ( $[M_{\text{sim}}]_{\text{LF}}$ ), photometric errors are simulated following the methodology outlined in Patapis et al. (2022).<sup>14</sup> Photometric errors of the same nature as resampling noise are therefore modeled by a sine function:

$$\epsilon(\lambda, x, y) = \sqrt{2}\sigma_{\text{ph-err}} \sin(2\pi f(x, y)\lambda), \quad (\text{E.1})$$

and the stellar modulation function with simulated photometric errors is given by

$$M_{\text{ph-err}}(\lambda, x, y) = [M_{\text{sim}}(\lambda, x, y)]_{\text{LF}} \times (1 + \epsilon(\lambda, x, y)), \quad (\text{E.2})$$

where  $\sigma_{\text{ph-err}}$  is the photometric calibration error and  $f(x, y)$  is a frequency randomly drawn for each spaxel. The probability density function of aliasing artifact frequencies is presumed to be dictated by the mean PSD (normalized) of the stellar modulation function,  $M_{\text{sim}}$  (with all corrections). This way, when the simulated photometric errors match the amplitude of those induced by aliasing artifacts in the data (measured at 7% by Law et al. 2023), both the mean PSD of the modulations and its spatial variation are recovered (see Fig. E.1). This suggests that the employed model for aliasing artifacts is suitable and adequate to recover the structure and variations in the modulations in the simulated data.

The contrast arising from systematics induced by these simulated photometric errors is computed using Eq. 13 with  $M_{\text{ph-err}}$  for each  $\sigma_{\text{ph-err}}$  and depicted in Fig. E.2. Likewise, when these simulated photometric errors match the amplitude observed in the data due to aliasing artifacts, the contrast obtained with  $M_{\text{sim}}$  (with all corrections, i.e., solid black line in Fig. 13) is reproduced.

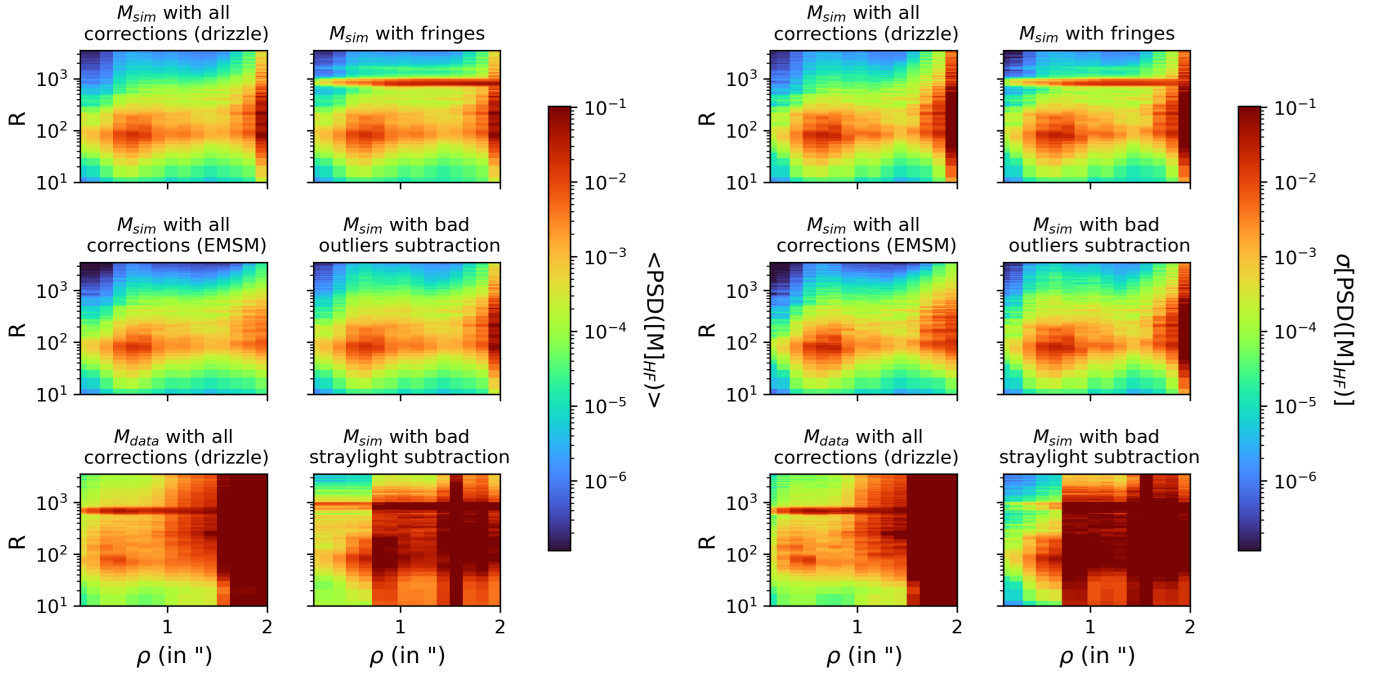
### E.2. Wavelength calibration errors

Simultaneously, it is crucial to assess the potential influence of wavelength calibration errors on noise. To achieve this, Doppler shifts were applied for each spaxel to the stellar modulation function without systematic high-frequency modulation (including the stellar spectrum):

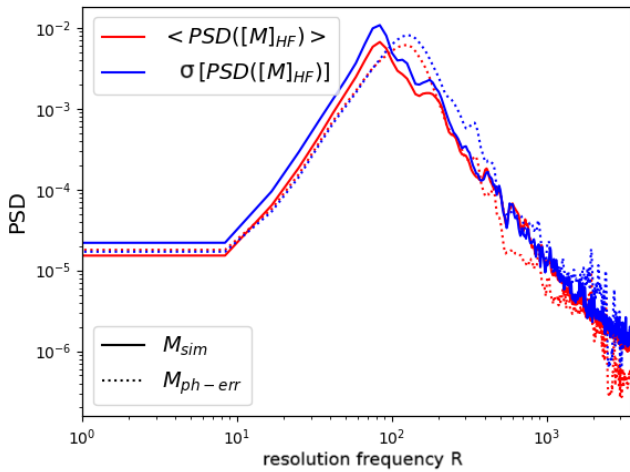
$$M_{\lambda\text{-err}}(\lambda, x, y)S_*(\lambda) = [M_{\text{sim}}(\lambda + \Delta\lambda(x, y), x, y)]_{\text{LF}}S_*(\lambda + \Delta\lambda(x, y)), \quad (\text{E.3})$$

where the offsets,  $\Delta\lambda(x, y)$ , were drawn randomly according to a normal distribution centered at 0 and with a standard deviation of  $\sigma_{\lambda\text{-err}}$ . As previously, the contrast due to systematics induced by these simulated wavelength errors was calculated according to Eq. 13 for each  $\sigma_{\lambda\text{-err}}$  and plot in Fig. E.3. It emerges that a calibration error of 100–1000 km/s would be required to explain the systematic noise observed with  $M_{\text{sim}}$  (with all corrections, i.e., solid black line in Fig. 13). In other words, when the simulated error matches the expected FLT-6 wavelength solution's amplitude error (few kilometers per second), the contrast is underestimated by a factor of roughly 30. Especially as the errors were simulated under extreme conditions: assuming no spatial correlation in the errors ( $\Delta\lambda(x, y)$ ), thus maximizing the systematic effect and the induced contrast. In fact, it is highly probable that there exists covariance in the wavelength error

<sup>14</sup> In that study, photometric and wavelength calibration errors are simulated to assess their impact on correlation. A similar approach is adopted in Sect. 5.4, but with the formalism detailed here and revised assumptions.

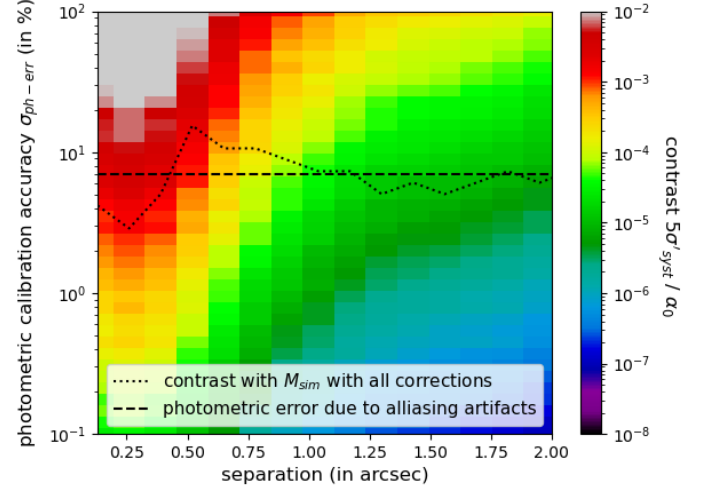


**Fig. D.1.** Left: Average of PSDs for each separation of the high-filtered stellar modulation functions on 1SHORT with  $R_c = 100$ . Right: Spatial variations in PSDs for each separation of the high-filtered stellar modulation functions (same as left panel but with standard deviation). The simulated stellar modulation functions  $M_{sim}$  were estimated by injecting a BT-NextGen stellar spectrum at 6000K into MIRISim. The stellar modulation functions at each point are renormalized so as not to have the flux dependency of the PSF (which would otherwise only show a PSD concentrated at  $\rho \approx 0$ ).  $M_{data}$  was estimated from HD159222 data. Edge effects, arising from both the detector and the cube reconstruction, begin to manifest at around 2 arcseconds.



**Fig. E.1.** Comparison of PSDs of stellar modulation functions when the amplitude of simulated photometric errors is equivalent to that of aliasing artifacts in the data. The modulation amplitude of the resampling noise on 1SHORT measured by Law et al. (2023) implies an error of  $\sigma_{ph-err} \approx 10\% / \sqrt{2} \approx 7\%$ . The mean and standard deviation are calculated all over the cubes.

between adjacent spaxels. For instance, assuming a covariance radius of 4 spaxels (deduced from Fig. C.2, since a plateau is reached after 4 pixels), the contrast is now underestimated by approximately a factor of 200. Hence, the impact of wavelength calibration errors on noise seems negligible. This also suggests that if only wavelength calibration errors (on the order of a few kilometers per second) were present as systematic noise, a contrast of  $10^{-4}$  could be achieved at 0.5'' and  $10^{-6}$  at 1.0'' for a planet at 1000 K.

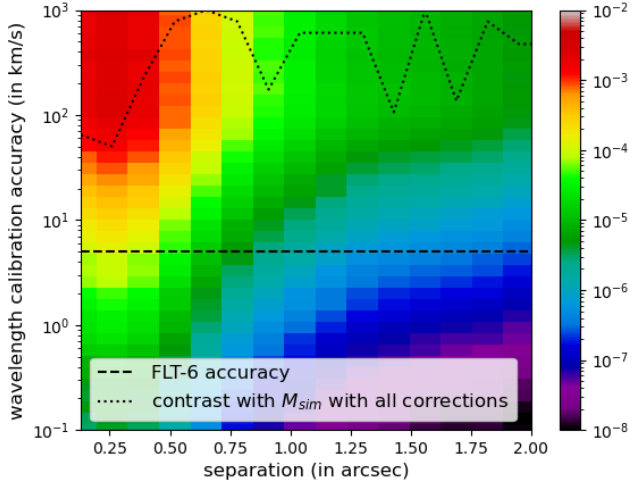


**Fig. E.2.** Simulated effect of photometric calibration errors of the same nature as aliasing artifacts on systematic noise on 1SHORT with a BT-Settl template at 1000 K and  $R_c = 100$ .

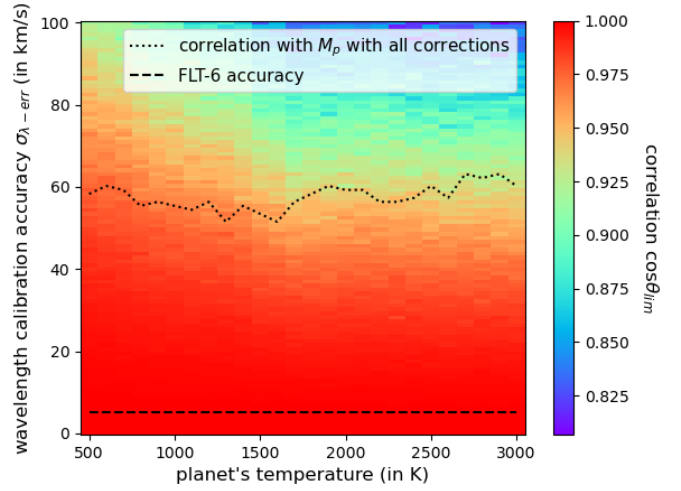
## Appendix F: Simulated systematic effects on correlation

### F.1. Photometric calibration errors

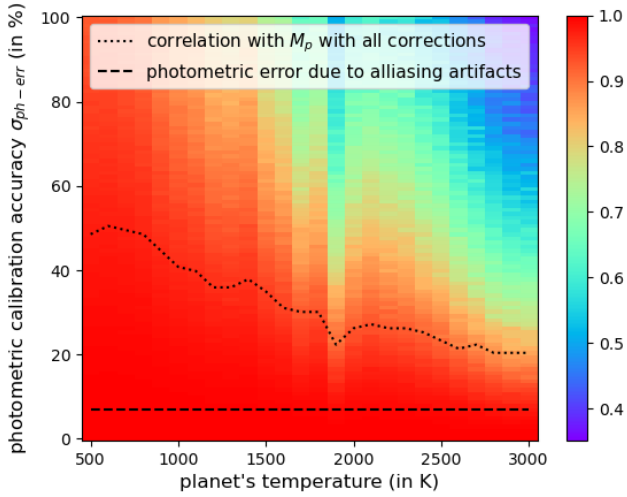
The idea is to replace  $M_p$  by  $M_{ph-err}$  in Eq. 15 and estimate the induced correlation drop for various photometric error levels  $\sigma_{ph-err}$  (see Fig. F.1).



**Fig. E.3.** Simulated effect of wavelength calibration errors on systematic noise on 1SHORT with a BT-Settl template at 1000 K and  $R_c = 100$ . The current wavelength solution (FLT-6) offers a calibration accuracy of a few kilometers per second in channels 1 and 2.



**Fig. F.2.** Simulated effect of wavelength calibration errors on correlation on 1SHORT with  $R_c = 100$ . The current wavelength solution (FLT-6) offers a calibration accuracy of a few kilometers per second in channels 1 and 2.



**Fig. F.1.** Simulated effect of photometric calibration errors of the same nature as aliasing artifacts on correlation on 1SHORT with  $R_c = 100$ . Photometric error due to resampling noise on 1SHORT:  $\sigma_{\text{ph-err}} \approx 7\%$ .

### F.2. Wavelength calibration errors

In the same way, replacing  $M_p$  with  $M_{\lambda\text{-err}}$  in Eq. 15 allows one to quantify the impact of wavelength calibration errors on the correlation (see Fig. F.2).