



# Direct emission spectroscopy of exoplanets with the medium resolution imaging spectrometer on board JWST MIRI

## I. Molecular mapping and sensitivity to instrumental effects

P. Patapis<sup>1</sup> , E. Nasedkin<sup>2</sup>, G. Cugno<sup>1</sup> , A. M. Glauser<sup>1</sup>, I. Argyriou<sup>3</sup>, N. P. Whiteford<sup>4,5</sup>, P. Mollière<sup>2</sup>, A. Glasse<sup>6</sup>, and S. P. Quanz<sup>1</sup>

<sup>1</sup> Institute for Particle Physics and Astrophysics, ETH Zurich, Wolfgang-Pauli-Str 27, 8093 Zurich, Switzerland  
e-mail: polychronis.patapis@phys.ethz.ch

<sup>2</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

<sup>3</sup> Instituut voor Sterrenkunde, KU Leuven, Celestijnenlaan 200D, bus-2410, 3000 Leuven, Belgium

<sup>4</sup> Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK

<sup>5</sup> Centre for Exoplanet Science, University of Edinburgh, Edinburgh, UK

<sup>6</sup> UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

Received 29 June 2021 / Accepted 25 October 2021

### ABSTRACT

**Context.** The Medium Resolution Spectrometer (MRS) of the Mid-Infrared Instrument (MIRI) on board the *James Webb* Space Telescope (JWST) will give access to mid-infrared (mid-IR) spectra (5–28 microns) while retaining spatial information. With the unparalleled sensitivity of JWST and the MIRI detectors, the MRS has the potential to revolutionise our understanding of giant exoplanet atmospheres.

**Aims.** Molecular mapping is a promising detection and characterisation technique used to study the spectra of directly imaged exoplanets. We aim to examine the feasibility and application of this technique to MRS observations.

**Methods.** We used the instrument simulator MIRISIM to create mock observations of resolved star and exoplanet systems. As an input for the simulator, we used stellar and planet parameters from literature, with the planet spectrum being modelled with the radiative transfer code petitRADTRANS. After processing the raw data with the JWST pipeline, we high pass filter the data to account for the stellar point spread function, and used a forward modelling approach to detect the companions and constrain the chemical composition of their atmospheres through their molecular signatures.

**Results.** We identified limiting factors in spectroscopic characterisation of directly imaged exoplanets with the MRS and simulated observations of two representative systems, HR8799 and GJ504. In both systems, we could detect the presence of multiple molecules that were present in the input model of their atmospheres. We used two different approaches with single molecule forward models, used in literature, that are sensitive to detecting mainly H<sub>2</sub>O, CO, CH<sub>4</sub>, and NH<sub>3</sub>, and a log-likelihood ratio test that uses full atmosphere forward models and is sensitive to a larger number of less dominant molecular species.

**Conclusions.** We show that the MIRI MRS can be used to characterise widely separated giant exoplanets in the mid-IR using molecular mapping. Such observations would provide invaluable information for the chemical composition of the atmosphere, complementing other JWST observing modes, as well as ground-based observations.

**Key words.** techniques: imaging spectroscopy – planets and satellites: gaseous planets – planets and satellites: atmospheres – infrared: planetary systems – methods: data analysis

## 1. Introduction

The past decade has proved revolutionary for exoplanet science with thousands of exoplanets being detected spanning a large range of masses, temperatures, and separations. While detection surveys are still ongoing with missions such as the Transiting Exoplanet Survey Satellite (TESS, Ricker et al. 2015) and *Gaia* (Gaia Collaboration 2016; Perryman et al. 2014), detailed characterisation of exoplanets with spectroscopic observations allows one to probe the properties of their atmospheres, as well as to have crucial links to the theory of exoplanet formation.

The small subset of exoplanets discovered with direct imaging offers a very distinct sample for the population of giant gas planets. These objects are mostly very young and hot (Bowler 2016), mainly due to observational bias; direct imaging

is sensitive to the thermal emission of these still contracting planets, which are all detected in the near-infrared (near-IR). The young age of these planets makes them the best targets for testing formation theories amongst the currently observed exoplanet population (e.g. Spiegel & Burrows 2012). By comparing radiation of giant planets with evolutionary track predictions for hot- and cold-start models, one can aim to reconstruct the formation pathway that the planet underwent (e.g. core accretion, gravitational instability; Pollack et al. 1996; Boss 1997) and its migration history (Mordasini et al. 2016).

Furthermore, emission spectroscopy enables the characterisation of their atmosphere properties such as the pressure-temperature profile (P-T profile), molecular composition, and cloud properties (Madhusudhan 2019). This provides insight into the early atmosphere evolution of gas giants as they cool down

over time (Bowler 2016). By carefully constraining elemental abundance ratios such as carbon-to-oxygen ratio (C/O), nitrogen-to-oxygen ratio (N/O), and sulphur-to-oxygen ratio, it is possible to directly link current atmospheric properties to formation, as the abundance ratios of the proto-planetary disk at the location where the planet formed is encoded into its atmospheric chemical properties (Öberg et al. 2011; Piso et al. 2016; Eistrup et al. 2016; Turrini et al. 2021). The history and evolution of forming planets as they interact with their natal environments can influence their chemical composition as well, by the amount of solids or gas that they accreted, or whether they migrated from their birth place to the observed location (Turrini et al. 2021; Öberg & Bergin 2021). For example one of the current theories behind the ammonia enrichment of Jupiter is based on formation beyond the N<sub>2</sub> ice-line (Öberg & Wordsworth 2019).

To date, all spectroscopic data of directly imaged exoplanets has been restricted to near-IR spectral bands (0.95–3.3  $\mu\text{m}$ , e.g. Konopacky et al. 2013; Barman et al. 2015; GRAVITY Collaboration 2017) from ground-based telescopes. With these measurements, the mass and radius cannot be empirically measured and are instead inferred from evolutionary and atmospheric models under certain assumptions such as hot and cold start (Bowler 2016). Some exceptions based on the combination of direct imaging and stellar astrometry ( $\beta$  Pic b, HR8799 Snellen & Brown 2018; Brandt et al. 2021) and on stability requirement for the planetary orbits (HR8799 system, Wang et al. 2018), offer important constraints for the models. The inference of other atmospheric properties such as temperature, surface gravity, clouds, and composition relies on the enhancing the spectral coverage and resolution, as well as improving atmospheric models that are fitted using Bayesian methods (e.g. Samland et al. 2017; Lavie et al. 2017; Wang et al. 2018, 2020, 2021; Nowak et al. 2020; Stolker et al. 2020a; Mollière et al. 2020). The lack of data for directly imaged exoplanets beyond 5  $\mu\text{m}$  (Petit dit de la Roche et al. 2019, 2021; Wagner et al. 2021) is mainly due to the dominating thermal background of earth’s atmosphere. Mid-IR observations would provide valuable and complementary information to near-IR data to address model degeneracy in the effective temperature estimation, impact of clouds and sensitivity to different molecules (Madhusudhan 2019; Bozza et al. 2018) and would give access to colder and older objects that are not detectable in the near-IR. The *James Webb* Space Telescope (JWST, Gardner et al. 2006) will help revolutionise the characterisation of exoplanet atmospheres, followed soon by the next generation of extremely large ground-based telescopes (ELTs).

The Mid-InfraRed Instrument (MIRI, Rieke et al. 2015b; Wright et al. 2015) on JWST is designed to serve many science cases, including exoplanet atmosphere characterisation. MIRI includes three observing modes, an imaging mode including coronagraphic filters, a medium resolution integral field spectrometer and a low resolution spectrometer. The Low Resolution Spectrometer (LRS, Kendrew et al. 2015) will mainly target transiting exoplanets with a resolving power of  $R \sim 50\text{--}100$  covering a wavelength range of 5–12  $\mu\text{m}$ . The MIRI Coronagraphic Imaging (Boccaletti et al. 2015) offers three four quadrant phase mask (FQPM) coronagraphs at 10, 11, 15  $\mu\text{m}$  and a Lyot coronagraph at 23  $\mu\text{m}$ , for imaging gas giant exoplanets in these filters. Although the usual angular differential imaging (ADI, Marois et al. 2008) won’t be as powerful as it has been for other facilities due to the small available telescope roll angle, the improved image stability and sensitivity compared to ground-based observations will provide the necessary contrast to observe most known objects, yielding 3 flux points from the FQPM coronagraphs, which can be used to constrain atmospheric models

(Danielski et al. 2018). The Medium Resolution Spectrometer (MRS, Wells et al. 2015) that is the focus of this work, will be described in detail in Sect. 2. It was not originally designed for exoplanet atmosphere characterisation, but rather for studying star formation and proto-planetary disks, as well as extra-galactic observations (Rieke et al. 2015b).

As the focus shifts to characterisation of exoplanets, there has been an increasing amount of interest by the exoplanet community in using integral field spectrometers (IFS) in the near-IR. Three of the current dedicated high contrast imaging instruments – SPHERE IFS (Beuzit et al. 2019) at the Very Large Telescope (VLT), GPI (Macintosh et al. 2014) at the Gemini Observatory and ScEXAO CHARIS (Currie et al. 2019) at the Subaru Telescope – have implemented such modes, combining extreme adaptive optics and coronagraphy with a spectrometer. These instruments have only been able to achieve a spectral resolution of  $R \sim 50\text{--}100$ , albeit at a high contrast. On the other hand, instruments such as VLT/SINFONI (Eisenhauer et al. 2003,  $R \sim 5000$ ), Keck/OSIRIS (Larkin et al. 2006,  $R \sim 4000$ ) and VLT/MUSE (Bacon et al. 2010,  $R \sim 3000$ ) have produced very interesting science output while lacking optimisation for direct imaging of planets. IFS instruments have allowed the detection of molecules in the atmospheres of HR8799 b/c (Konopacky et al. 2013; Barman et al. 2015; Petit dit de la Roche et al. 2018),  $\beta$  Pictoris b (Hoeijmakers et al. 2018; Rameau et al. 2021), kappa And b (Todorov et al. 2016; Wilcomb et al. 2020) and HIP65426 b (Petrus et al. 2021), measurements of the radial velocity of HR8799 b and c (Ruffio et al. 2019) and the investigation of physical and chemical properties of the forming planet PDS70 b and its surroundings (Cugno et al. 2021). This is made possible by taking advantage of the improved spectral resolution to disentangle the fundamentally different planetary and stellar signals, using cross-correlation techniques with forward models. In the near future, VLT/ERIS (Davies et al. 2018) and Keck/KPIC (Mawet et al. 2016) will provide upgraded spatially resolved spectroscopy on ground-based telescopes, and together with NIRSpec (Bagnasco et al. 2007; Ygouf et al. 2017) on board JWST, will improve the characterisation of directly imaged exoplanets in the near-IR. Despite the large variety of existing and future near-IR instruments, the only planned mid-IR IFS are the MIRI-MRS slated to launch in 2021 and ELT/METIS (Brandl et al. 2021) that is expected to see first light in 2028.

In this work we investigate the potential of the MRS to detect molecules in the atmospheres of young giant exoplanets. In Sect. 2 we introduce the instrumental specifications of the MRS relevant to observing faint signals and the tools to create simulated data of our targets. We then describe the cross-correlation data reduction steps for medium resolution IFS data in Sect. 3. Finally in Sect. 4 we demonstrate how the MRS can provide insights into the chemistry of directly imaged planet atmospheres, both through molecular mapping, and a likelihood based framework for the detection of various molecular species, and discuss the results in Sect. 5.

## 2. Observing self-luminous companions with the MRS

### 2.1. MIRI medium resolution spectrometer

The Medium Resolution Spectrometer (MRS, Wells et al. 2015) on MIRI covers a wavelength range of 4.8–28.5  $\mu\text{m}$  with a resolving power of 4000–1500, decreasing with higher wavelength. It is divided into four overlapping wavelength channels (1–4) with two Si:As IBC detectors for channels 1/2 and 3/4

respectively, that are read out non-destructively (Rieke et al. 2015a). Additionally each channel is split into 3 spectral sub-bands (A, B, C) that all need to be observed in order to cover the full wavelength range. Since all channels can be exposed simultaneously the full wavelength coverage of the MRS requires just 3 exposures, one for each sub-band. Being a IFS, the MRS retains both spatial coordinates with the end data product being a data cube where each pixel is a spectrum, denoted as a spaxel. The spatial resolution, field of view, and wavelength coverage of each sub-band can be seen in Table 1. These specifications provide some advantages over ground-based instruments of the same type, with a few key factors making the MRS unique.

First, the low thermal background of JWST/MIRI will provide unprecedented sensitivity for faint sources in the 5–28  $\mu\text{m}$  range (Glasse et al. 2015), compared to the limited sensitivity of ground-based exposures due to atmosphere’s thermal background (e.g. VLT/VISIR, VLT/NEAR, Petit dit de la Roche et al. 2019, 2021; Pathak et al. 2021). Second, considering only the first 3 channels up to 18  $\mu\text{m}$ , since channel 4 suffers for a great drop in sensitivity, the continuous wavelength coverage is almost 10 times larger than current near-IR instruments, which will greatly benefit fitting atmosphere models and finding features of multiple molecular species (see Fig. 1). Lacking any mid-IR spectra, molecular features of giant planet atmospheres are currently not constrained. Thus, the MRS will add novel invaluable observational evidence to further improve and constrain existing atmospheric models. Additionally with the cross-correlation technique applied to IFS data we probe another aspect of the data, namely narrow molecular features from many different molecules (Woitke et al. 2018). Third, the contrast of exoplanets with respect to their host star is much more favourable in the mid-IR, with the stellar spectrum decreasing rapidly with wavelength ( $\propto \lambda^{-4}$ ). At the same time, the features of typical stellar spectra at these wavelengths are originating from atomic absorption lines and will not correlate with the exoplanet atmosphere signature mostly dominated by molecules (Bozza et al. 2018). Finally, the point spread function (PSF) is expected to be much more stable in comparison to the near-IR, with none of the atmospheric turbulence seen in ground-based observing and well-characterised diffractive broadening dominating the spatial profile, compared to the alignment errors seen in the multi-element optical train which are more important at shorter wavelengths (Lightsey et al. 2014). The gain in Strehl ratio in space will be an advantage for disentangling and modelling stellar and planet signal in our data.

Nonetheless, the complexity of the MRS requires precise calibration and pre-processing of the raw data. The Integral Field Units are based on a slicing optic (in contrast to a lens-let array implemented for instance in GPI and OSIRIS), which results in a different sampling for each spatial coordinate (Glauser et al. 2010). The optical design of the spectrometer introduces distortion of the dispersed light on the detector resulting in a non-orthogonal projection on it (Wells et al. 2015). The most prominent source of systematic error for the MRS is fringing caused by interference within the layers of the detector. It is expected to be more significant in the mid-IR as the wavelength of the light is comparable to the thickness of the detector layers and can cause up to 20% flux variation as a function of wavelength (Argyriou et al. 2020). Additionally, a straylight component (Wells et al. 2015), adds to the photometric uncertainty. Observations of point sources are always performed using dithered exposures in order to improve the spatial sampling, as the spatial resolution is not Nyquist sampled for the short wavelength sub-bands (Glauser et al. 2010). Following multiple steps of the calibration pipeline (Labiano et al. 2016) the data

**Table 1.** MRS instrument specifications.

Band	FoV ["] <sup>(a)</sup>	Spatial res ["] <sup>(a)</sup>	$\lambda$ [ $\mu\text{m}$ ] <sup>(b)</sup>	$R$ <sup>(b)</sup>
1A	$3.2 \times 3.7$	0.196	4.88–5.77	3500
1B			5.64–6.67	
1C			6.50–7.70	
2A	$4.0 \times 4.8$	0.196	7.47–8.83	3200
2B			8.63–10.19	
2C			9.96–11.77	
3A	$5.5 \times 6.2$	0.245	11.49–13.55	2500
3B			13.28–15.66	
3C			15.34–18.09	
4A	$6.9 \times 7.9$	0.273	17.60–21.00	1700
4B			20.51–24.48	
4C			23.92–28.55	

**Notes.** FoV is the instrumental field of view,  $\lambda$  denotes the spectral coverage, and  $R$  refers to the average resolving power of the channel.

**References.** <sup>(a)</sup>Wells et al. (2015). <sup>(b)</sup>Labiano et al. (2021).

are calibrated (described in detail in Sect. 2.6) and the detector images are reconstructed into a data cube, where an image is produced for each wavelength bin (Wells et al. 2015; Labiano et al. 2021). Residual systematic errors in each step of the calibration pipeline pose a significant obstacle to observations of very faint sources, such as exoplanets around a bright star, and careful correction of these residual effects might be critical to the success of the MRS as an exoplanet characterisation instrument.

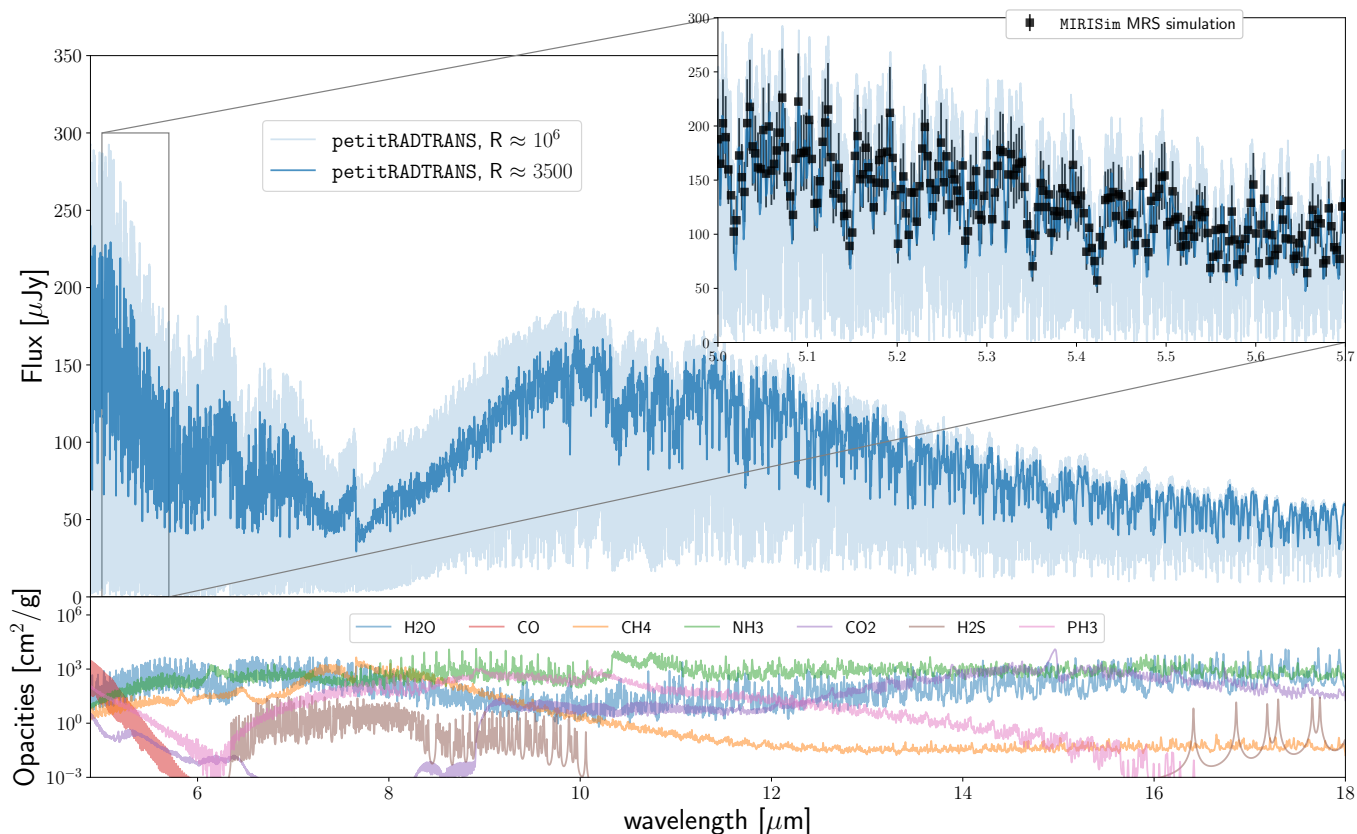
## 2.2. Observational limits

The design and instrumental effects of the MRS place several constraints on the sample of giant exoplanets that can be characterised with it. In this section we present the main instrumental effects that dictate which planets are observable with the MRS: (i) bright source limit, (ii) spatial resolution, (iii) MRS sensitivity, and (iv) additional observing considerations related to the exoplanet system geometry and alignment with respect to the used dither pattern.

(i) The MRS is expected to saturate on sources brighter than 4–5 Jy (Glasse et al. 2015). If the star is contained in the field of view of the scene that will be observed, the lack of a coronagraph means there is an upper limit to the length of integration that can be used without saturating the detector. In principle, the non-destructive readout of the detectors would allow for saturated pixels<sup>1</sup>, with the pipeline using the partial ramp to fit the slope in post-processing. However, saturation for the MIRI detectors causes long lasting (up to several hours) persistence and is responsible for a series of effects that can affect the region around the saturated pixels, due to cross-talk between pixels (Rieke et al. 2015a). Reducing the length of the integration impacts the signal-to-noise ratio (S/N) of the planet and will result in longer exposure time (consisting of multiple integrations) being required.

(ii) The diffraction limited images delivered by the JWST have a diameter of the first dark Airy ring equal to 1.0 arcseconds at a wavelength of 10  $\mu\text{m}$ , making it difficult to resolve planets with angular separations less than 0.5 arcseconds from their parent star.

<sup>1</sup> For Cycle 1 observations saturation of the detectors is not permitted.



**Fig. 1.** *Top:* atmospheric model of HR8799 b calculated with `petitRADTRANS` and simulated MRS data using `MIRISIM` with realistic observation parameters, including all available noise sources. After the whole simulation chain and pipeline calibration steps we recover the input flux (see inset) by subtracting the stellar PSF with a simulation of HR8799 A without the planets. The error bars indicate the  $3\sigma$  estimates from apertures with no planet signal at the same separation from the star as HR8799 b. *Bottom:* opacity of different molecules as a function of wavelength in the mid-IR. Many molecules provide modulation in their opacity which will translate to detectable molecular features in their spectra.

(iii) MIRI’s sensitivity is unprecedented for the mid-IR. However, we note that the MRS’ high spectral resolution makes it less sensitive for detection of a planet’s continuum spectrum, compared to the MIRI-LRS or imager, with the MRS limited to approximately  $35 \mu\text{Jy}$  per spectral pixel (or equivalently a 18.2 K-band magnitude) for a  $5\sigma$  signal in 10 000 s (Glasse et al. 2015).

(iv) There are two aspects of observing with the MRS that need special attention when observing spatially resolved planetary systems. First, the small Field of View (FoV – just  $3.2 \times 3.7$  arcseconds in the short wavelength channel) makes the sky alignment of the instrument with respect to the exoplanetary system important. Especially when multiple companions are present, sub-optimal positioning of the field and choice of dither pattern might influence the signal of the planets. We built a visibility tool for the MRS that can help plan observations and choose a dither pattern illustrated in Fig. A.1. Secondly, the detectors of MIRI are susceptible to persistence when bright sources are observed (Rieke et al. 2015a). Placing the planet at a position in the field where the star was located in a previous exposure will affect its appearance due to the additional latent image. Therefore the dither pattern choice should also consider this effect when planning observations.

From the Extrasolar Planets Encyclopaedia<sup>2</sup> we collected all sub-stellar companions that fit the constraints outlined above. When limiting the sample to objects with reported mass in the

planetary mass regime, the known objects that are observable with the MRS are shown in Table B.1. Two notable directly imaged exoplanet systems are missing: PDS70 b,c (Keppler et al. 2018; Haffert et al. 2019) and 51 Eridani b (Macintosh et al. 2015). These are very interesting objects, the former being a forming planetary system whose companions are potentially surrounded by a circumplanetary disk (Stolker et al. 2020a; Cugno et al. 2021; Benisty et al. 2021), and the latter one of the coolest imaged exoplanets with a temperature of  $\sim 700$  K (Rajan et al. 2017; Samland et al. 2017). The PDS70 planets are too close to their host star to be spatially resolved, while 51 Eri A is over the brightness limit of the MRS, saturating even with the minimum length of integration. Due to the non-destructive readout of the MRS detectors 51 Eri b may become feasible when the technical problems of calibrating partially saturated images are well developed.

### 2.3. Two case studies for the MRS

For the rest of this work we use two systems as a test-bed to illustrate the capability of the MRS to characterise exoplanets. These are the HR8799 system (Marois et al. 2008, 2010) that hosts 4 known planets, and GJ504 b (Kuzuhara et al. 2013; Janson et al. 2013). The two targets span a pivotal temperature regime that provides insights into the chemical processes of gas giant exoplanets, with one hosting hot planets and the other a cold planet. The HR8799 planets have been extensively studied in the near-IR and are all estimated to have an effective temperature between

<sup>2</sup> <http://exoplanet.eu>

1000 and 1200 K (Marois et al. 2008; Bonnefoy et al. 2016; Mollière et al. 2020). One of the characteristics of this temperature regime for giant exoplanets is the change in the dominant carbon-bearing molecule in the observable photosphere, from CO to CH<sub>4</sub> for cooling temperatures (Madhusudhan et al. 2016). While Barman et al. (2015) detected CH<sub>4</sub> in HR8799 b, more recently Petit dit de la Roche et al. (2018) and Ruffio et al. (2021) could not confirm this detection. Such open questions could be tackled with the extended wavelength coverage and sensitivity achieved by the MRS, placing limits on molecular abundances. The second system, GJ504 b has an effective temperature of ~550 K (150 K cooler than 51 Eri b), and was likely formed through gravitational instability (Bonnefoy et al. 2018). Due to its low temperature it is expected to harbour both CH<sub>4</sub> (that has already been detected by Janson et al. 2013) and NH<sub>3</sub> observable signatures in its SED, making GJ504 b a prime target to investigate the atmosphere physics of cooler gas planets.

#### 2.4. Emission spectra model

For calculating all emission spectra and forward models of the giant gas planets we use the 1D radiative transfer code `petitRADTRANS` (Mollière et al. 2019; Mollière et al. 2020). With a user friendly and fast python implementation<sup>3</sup>, this code takes as an input the thermal structure profile of the atmosphere (P–T profile), its gravity, and a list of continuum and line opacity sources together with the abundance of H, He, and the list of molecules included. It includes parameters for modelling clouds which we did not use in the main analysis. Clouds are expected to have distinguishable spectral features for mid-IR wavelengths at this spectral resolution (Wakeford & Sing 2015; Powell et al. 2019; Gao et al. 2020), where broad features will be filtered out during the continuum removal, and higher frequency features would overlap other molecular features and could potentially be detected with cross-correlation. For the purpose of this work we consider clouds to overall reduce the contrast of molecular features, negatively impacting the detection of molecules.

One of the trade offs of this code is that the calculated emission spectra are not necessarily physical. In this work we assume a constant abundance of each molecule throughout the whole atmosphere altitude, for which the mass fraction value is arbitrarily set. To make sure that our models are at least plausible and representative of the atmospheres of known directly imaged planets we use the following inputs for the (i) p–T profile and the (ii) chemical composition of the atmosphere.

(i) The p–T profile is taken from a grid of self-consistent models originally published in Samland et al. (2017). This grid was produced with `petitCODE` (Mollière et al. 2015) and is parameterised by the effective temperature ( $T_{\text{eff}}$ ), gravity ( $\log(g)$ ), and metallicity (Fe/H).

(ii) The authors of `petitRADTRANS` offer a very useful tool for determining the equilibrium chemistry of an atmosphere, wrapped in a python package that is called `poor_mans_disequ_chem`<sup>4</sup>. It is parametrised by effective temperature, gravity, C/O, and Fe/H, and even includes a simple quenching mechanism (Mollière et al. 2017). We use this tool in order to obtain a better physically grounded abundance value of each species in our models, calculating the abundance distribution of the different species as a function of altitude and selecting the abundance value for a given molecule at the pressure of  $10^{-2}$

**Table 2.** Atmosphere parameters used in our `petitRADTRANS` simulation.

Parameter	HR8799 b <sup>(a)</sup>	HR8799 c, d, e <sup>(b)</sup>	GJ504 b <sup>(c)</sup>
$R_{\text{planet}} [R_{\text{J}}]$	1.2	1.3	1.0
Distance [pc]	41	41	17.5
$\log(g)$ [cgs]	4.0	4.0	3.9
$T_{\text{eff}}$ [K]	1100	1200	550
Fe/H [dex] <sup>(*)</sup>	0.2	0.2	0.2
C/O [dex] <sup>(*)</sup>	0.4	0.4	0.4
Molecule	$\log(X_{\text{mol}})$		
H <sub>2</sub> O	-2.1	-2.2	-2.0
CO	-2.8	-2.3	<-10
CH <sub>4</sub>	-2.4	-2.7	-2.3
NH <sub>3</sub>	-5.0	-5.3	-2.9
H <sub>2</sub> S	-3.3	-3.3	-3.3
PH <sub>3</sub>	-5.0	-5.1	-8.0
CO <sub>2</sub>	-5.3	-5.0	<-10

**Notes.** <sup>(\*)</sup>The parameters of Fe/H and C/O were kept constant in order to have a more comparable atmosphere chemistry.

**References.** <sup>(a)</sup>Wang et al. (2018). <sup>(b)</sup>Wang et al. (2020). <sup>(c)</sup>Bonnefoy et al. (2018).

**Table 3.** Parameters of stellar spectra.

Parameter	HR8799 A <sup>(a)</sup>	GJ504 A <sup>(b)</sup>
$R_{\text{star}} [R_{\odot}]$	1.5	1.2
$\log(g)$ [cgs]	4.3	4.6
$T_{\text{eff}}$ [K]	7500	6300

**References.** <sup>(a)</sup>Gray & Kaye (1999). <sup>(b)</sup>Kuzuhara et al. (2013).

bar. We note that the abundances are still assumed to be vertically constant during our spectral modelling with the value found from the chemical equilibrium calculation above.

The list of parameters that are used in our `petitRADTRANS` models are shown in Table 2. We kept the values of Fe/H and C/O fixed for all the planets in order to have more comparable atmospheres. To motivate that the calculated model spectra are a good description of the planets we simulate, we used the `species` python package (Stolker et al. 2020b) to retrieve near-IR data of HR8799 b and GJ504 b and plot them in Fig. 2.

It should be noted that none of these models include the spin rate of the planet surface that would for higher resolution spectra affect the cross-correlation signal, seen as a broadening (Snellen et al. 2014). At the spectral resolution of the MRS this effect is negligible.

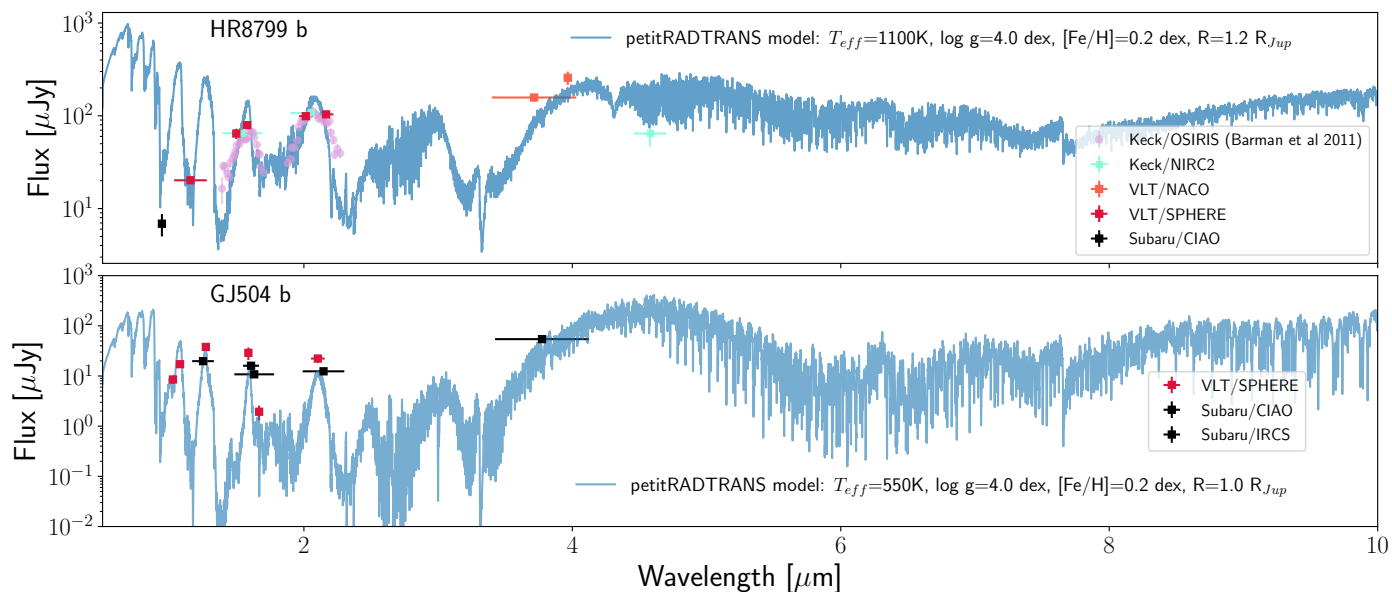
For the stellar models in our simulations we use the BT-NEXTGEN models (Allard et al. 1997) accessed through `species` with the parameters of each host star taken from literature (Marois et al. 2008; Kuzuhara et al. 2013), shown in Table 3.

#### 2.5. Data simulation – mock observations

We use the MIRI instrument SIMulator (MIRISIM, Klaassen et al. 2021, version 2.3) to simulate realistic observations of the MRS. MIRISIM is an end-to-end simulator that allows the user to define a custom astronomical scene and, given the observation parameters, it propagates the scene through a simulation where

<sup>3</sup> <https://petitradtrans.readthedocs.io/en/latest/>

<sup>4</sup> [https://petitradtrans.readthedocs.io/en/latest/content/notebooks/poor\\_man.html](https://petitradtrans.readthedocs.io/en/latest/content/notebooks/poor_man.html)



**Fig. 2.** Atmospheric model calculated with `petitRADTRANS` and archival near-IR data retrieved from species for HR8799 b (*top*) and GJ504 b (*bottom*). The models capture the observed spectral energy distribution and should provide a plausible emission spectrum of the planet atmosphere in the mid-IR.

many known detector and instrument effects are modelled based on MIRI data obtained during several ground test campaigns.

For the simulations of our data we define a point source in the MIRISIM scene for the star and each of the planets in the system. The location of the point sources is defined in RA and Dec with the centre of the MRS field as the zero-point. This is very convenient as the astrometry of exoplanets is often reported as  $\delta$ RA and  $\delta$ Dec with respect to the star, or as separation from the star and the position angle (PA) with respect to the north. Each point source is assigned a spectral energy distribution (SED) with an array of wavelengths ( $\mu\text{m}$ ) and fluxes ( $\mu\text{Jy}$ ).

The observation parameters were optimised with the JWST Exposure Time Calculator<sup>5</sup>. The input parameters include detector readout mode (FAST or SLOW), number of groups per integration (NGROUP – effectively the number of readouts per integration), integrations per exposure (NINT), and dither pattern. For the MRS and our observations of exoplanets we always use FAST mode, with each frame exposed for 2.755 s. NGROUPS is optimised to maximise the number of groups such that the star does not saturate the detector. We choose NINT as necessary to achieve the desired S/N for the given exoplanet, while the adopted dither pattern is defined as described in Appendix A.

We include all available instrumental and detector effects in our simulations, except cosmic rays, which have been shown to not behave as expected in MIRISIM simulations and were disabled. The detector effects include photon noise, read noise, bad pixels, dark current, gain and detector flatness effect, non-linearity of the ramp, detector fringing, and a simple latency effect (Klaassen et al. 2021). Instrumental effects include spatial and spectral distortion, PSF broadening in one of the spatial dimensions, and photon conversion efficiency, which includes the quantum efficiency of the detectors and the instrument optical chain transmission.

Using the procedure for creating MRS observations we simulated data for HR8799 and GJ504. For the HR8799 system we placed the star in the centre of the FoV and observed for 8 h

using NGROUPS = 22, NINT = 40, and the extended source 4 point dither pattern (Fig. A.1). GJ504A was offset such that the star falls just out of the FoV in Channel 1 and GJ504 b is centred in the field. This was essential since the star is over the bright source limit of the MRS in Channel 1. We used NGROUPS = 50, NINT = 10, and the point source 4 point dither pattern for a total exposure time of 4.5 h. In the inset of Fig. 1 we show that the simulation correctly reproduces the input flux of HR8799 b. The flux was retrieved by simulating an observation with the same parameters as the HR8799 system but only containing the star, and then subtracting it to isolate the planet signal. Using an aperture with a radius encompassing 80% of the PSF model we extracted the flux for each wavelength bin of the reconstructed cube.

## 2.6. JWST data reduction pipeline

MIRISIM delivers raw data in the form of FITS files exactly as the JWST data products that will be provided to observers on the Mikulski Archive for Space Telescopes (MAST). These detector raw data consist of one image for every frame, integration and exposure, in units of digital numbers (DN). DN represent the incident photons that were converted to electrons, read out by the detector, and stored in memory. In order to process these files and bring them to flux calibrated data cube form we use the JWST Data Reduction Pipeline<sup>6</sup>. A detailed description of the pipeline steps applied to MRS raw data can be found in Labiano et al. (2016).

In the first stage (`calwebb_detector1`) the up-the-ramp samples of the non-destructive detector readouts in DN are transformed into count rate (DN/s). Here various detector level effects from the MIRI detectors are mitigated, such as: first and last frame effects, non-linearity of the ramps, detection of cosmic rays (not applicable in our case), and dark current correction. In the end a linear function is fitted to each pixel ramps and the slope is saved as the count rate of the given pixel, along with a data quality and error value.

<sup>5</sup> <https://jwst.etc.stsci.edu>

<sup>6</sup> <https://jwst-pipeline.readthedocs.io/en/latest/>

Next, in the second stage of the calibration pipeline (`calwebb_spec2`) all instrument specific physical corrections are applied. The goal is to produce a flux calibrated detector image for each exposure. These include the background subtraction, world coordinate system assignment, fringe correction, and spectro-photometric calibration.

Finally in the last stage (`calwebb_spec3`) the different exposures for each dither position are reconstructed into a cube. This cube building is a critical step in the pipeline since it affects spatial and spectral resolution, introducing artefacts and correlated noise to the data (Liu et al. 2020). The current pipeline is using an exponentially decaying weight function to interpolate the flux of a given spaxel. The output of the JWST Data Reduction Pipeline consists of a data cube per MRS spectral sub-band.

### 3. Cross correlation technique

Cross-correlating a forward model of the spectrum with the IFU data has been very successful in isolating the exoplanet signal from the stellar PSF and the noise (Hoeijmakers et al. 2018; Ruffio et al. 2019). It is in essence a method of collapsing the data using weights given by a forward model, which results in boosting the S/N of spaxels where the spectra agree better with the forward model. This processing method is especially advantageous when high fidelity modelling, and subtraction, of the stellar PSF from the data down to the contrast level of the planet is not possible, as for example due to the lack of a coronagraph or instrument design.

Applying cross-correlation to observations of exoplanets in the mid-IR with a space telescope has very clear advantages compared to similar observations from ground-based IFS. Indeed, (i) the lack of atmospheric speckles, (ii) the absence of telluric lines (Hoeijmakers et al. 2018), and (iii) the simplicity of the stellar spectrum in the Planck tail of the black-body with only a few atomic absorption lines will result in a low frequency component of each spaxel (Ruffio et al. 2019), even in the presence of low order wavefront errors from the telescope. Conversely, the exoplanet emission spectrum should contain many molecular absorption features and lines, with the opacity of different molecules being more prevalent at different wavelengths as shown in the bottom panel of Fig. 1. These features will translate to high frequency modulation of the spectra that can be separated from the dominating stellar signal in the data as we show below, for star-to-planet contrast ratio down to  $\sim 10^{-4}$  that is considered challenging even for coronagraphic observations with MIRI (Danielski et al. 2018).

For the data processing leading up to the cross-correlation we used a high-pass filter (HPF) by following a simplified version of Hoeijmakers et al. (2018). The full processing used in this work consists of the following steps. The JWST pipeline was applied, producing a reconstructed cube with a given spatial and spectral sampling. For each spaxel the dominating stellar continuum was estimated by convolving the spaxel with a Gaussian and outliers were removed with sigma clipping of 10 standard deviations. This continuum was then subtracted from the spaxel, with only the high frequency component remaining in the data. The HPF spaxels were then normalised by dividing with their maximum absolute value. The forward model was convolved with a Gaussian to the average spectral resolution of the MRS for each sub-band and binned<sup>7</sup> to the wavelength coordinates of the data,

<sup>7</sup> The python package `spectres` (Carnall 2017) was used to bin a high resolution spectrum to a new wavelength grid.

and then high-pass filtered and normalised in the same manner as the data. Finally, each spaxel was cross correlated<sup>8</sup> with the forward model producing the cross-correlation function (CCF)

$$c_{av}[k] = \frac{\sum_n a[n+k]v[n]}{\sqrt{a^T a v^T v}},$$

where  $k$  is an index indicating the step of the CCF.

In Fig. 3 the processing is illustrated for a simulated data set where a random sample of spaxels was chosen (panel a) and their power spectral density (panel b), HPF spectrum (panel c) and cross-correlation with the forward model (panel d) is shown. The value for the Gaussian kernel standard deviation was chosen after varying the parameter and choosing the one producing the highest cross-correlation S/N (as defined in Sect. 4.1) when using the input model as the forward model. We expect this value to be a free parameter that would be optimised for each observation depending on the continuum of the stellar spectra and the PSF modulation as a function of wavelength. The spaxel containing the planet signal produces a peak in the cross-correlation while the rest of the spaxels indicate the noise floor from the various noise terms (Fig. 3d).

We note that the PSF of the star was not explicitly modelled but could be included in the forward model if there is an available stellar model, as for example shown in Ruffio et al. (2019). Additional methods of dealing with remaining systematic residuals such as principal component analysis (e.g. Cugno et al. 2021) were not applied as the simulated data do not provide such effects, which are difficult to predict before having access to real on-orbit data. After the last step in the data processing the final data product is a cube with the two spatial dimensions of the original cube, but with the spectral dimension now replaced by the values of the CCF for each velocity shift. The range of velocity shift for the CCF can be chosen but is in principle only limited by the wavelength range of the forward model.

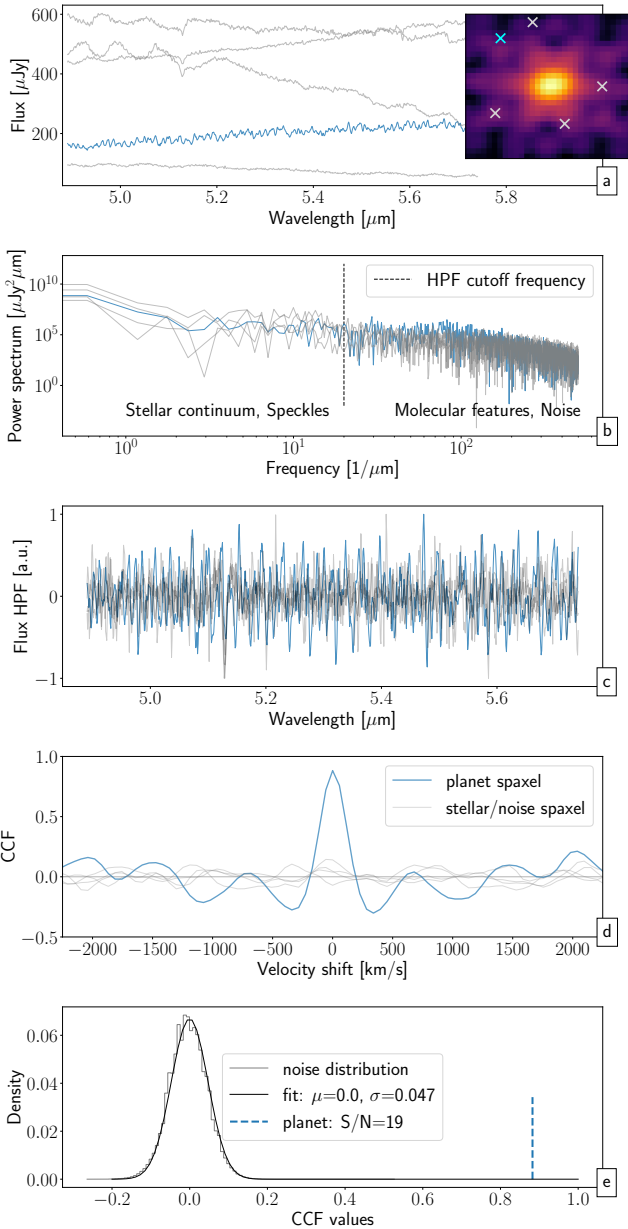
## 4. Analysis and results

### 4.1. Molecular mapping in the mid-IR

For the characterisation of these exoplanets with the MRS we started with the flux calibrated spectral cube and cross-correlated each spaxel with a template as described in Sect. 3. We used both the atmosphere model used as the input for the planet spectrum in the simulation, and molecular templates featuring the absorption signature of different dominating molecules such as H<sub>2</sub>O, CO, CH<sub>4</sub>, and NH<sub>3</sub>. These molecule template spectra were again calculated with `petitRADTRANS` by using the same P-T profile as for the input model but only considering the continuum contributors of H and He, together with the specific molecule opacity with a molecular mass fraction of  $10^{-2}$ . This spectrum may be nonphysical but it provides a template for the spectral features and lines that each molecule contributes to the emission spectrum of the exoplanet atmosphere. The continuum subtracted (as in Sect. 3) molecule templates with the P-T profile used for HR8799 b are shown in Fig. 4.

The S/N for the cross-correlation signal here is defined as in Petrus et al. (2021). A t-test of the CCF signal at the radial velocity and position of the planet is performed with respect to the noise distribution of the CCF, illustrated in the Fig. 3e. All pixels at a location not containing the planet signal and at radial velocity shifts larger than  $500 \text{ km s}^{-1}$  (equivalent to 10 pixels in

<sup>8</sup> The python function `numpy.correlate` was used to perform the one dimensional correlation between two arrays.



**Fig. 3.** Processing steps of MRS reconstructed cubes for the application of the cross-correlation technique. (a) Spectra at different locations in the FoV of sub-band 1A of the HR8799 observation were selected as shown in the inset (image in log scale). The spaxel containing planet signal (in blue) is not directly visible in the cube and is dominated by the stellar PSF. (b) Power spectrum of the selected spectra. By visualising the frequency domain we can see where the high pass filter operates and distinguishes continuum contribution dominated by the star and high frequency components that contain the molecular absorption lines and residual noise. (c) High pass filtered and normalised flux. (d) Signal of the cross-correlation function where the input spectrum was used as a forward model. The planet spaxel clearly correlates with the forward model with a peak value of 0.88 ( $S/N \sim 19$ ). (e) S/N calculation for the cross-correlation method. All noise spaxels were considered and a Gaussian was fitted to the distribution to estimate the standard deviation that is used to estimate the S/N for a given cross-correlation value of the planet spaxel.

the spectral direction) at the location of the planet were considered as noise samples. Before the calculation of the noise the auto-correlation of the forward model scaled to the signal of the CCF at the RV of the planet was removed was each spaxel, since

it is evident that it dominates the noise for high correlation values, as seen in Fig. 3d. The distribution of the noise sample is fitted with a Gaussian to estimate its mean and standard deviation. These are then used in the t-test to calculate the S/N as

$$S/N = \frac{(\mu_{\text{planet}} - \mu_{\text{noise}})}{\sigma_{\text{noise}}}, \quad (1)$$

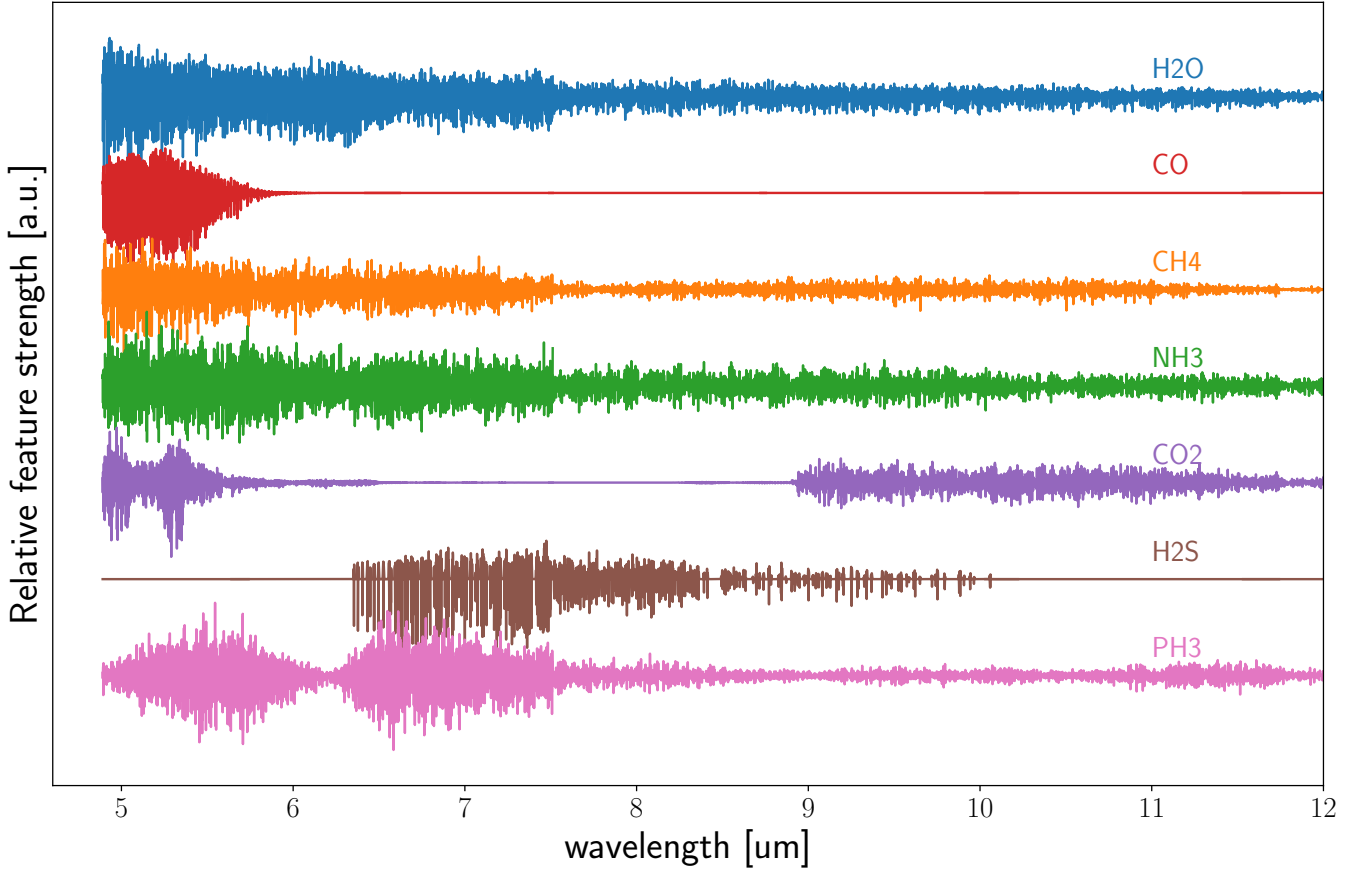
where  $\mu$  and  $\sigma$  are the mean and the standard deviation, respectively. The planet CCF mean is calculated from an aperture centred on the planet.

The results are summarised in Fig. 5 for HR8799 and Fig. 6 for GJ504. In Fig. 5 the top left panel a single frame from the spectral cube is shown with the location of the planets annotated with circles. Below, the cross-correlation with the forward model of the full atmosphere used as an input, reveals three of the four HR8799 planets: b, c, and d. The fourth planet HR8799 e is not detected since it is located too close to the star and its signal is lost under the noise limit. On the right, the four panels show the result of molecular mapping for the molecular species with the highest abundance.  $\text{H}_2\text{O}$  and  $\text{CO}$  with a log mass fraction of  $-2.3$  and  $-2.7$  are detected with S/N of 16.5 and 3.6 for the HR8799 b planet. These molecules had previously been confirmed in Barman et al. (2015) in the near-IR.  $\text{NH}_3$  and  $\text{CH}_4$  with log mass fraction of  $-5.0$  and  $-2.4$  are detected at an S/N of 4.9 and 6.8. Respectively for GJ504 b in Fig. 6, molecules with high abundance such as  $\text{H}_2\text{O}$ ,  $\text{NH}_3$ , and  $\text{CH}_4$  are significantly detected with an S/N of 19.7, 15.9, and 7.5. These results provide detection of different species in the atmosphere of a planet without any prior knowledge of specific parameters for the planet atmosphere other than the line list that produced this molecule opacity in `petitRADTRANS`. Although the P–T profile does affect the line contrast, the molecule template spectra are less sensitive on the specific choice we make due to the fact that we normalise it before cross-correlating with the data. It is important though to produce a template that highlights the features of the spectrum that the cross-correlation might pick up. This provides direct detection of molecules that are present in the atmosphere and can inform the next steps of the analysis.

#### 4.2. Likelihood ratio test for molecule detection

While using a single molecule forward model to infer the presence of a molecule works for the dominant species, it might introduce a bias by not accounting for the impact of other species on the underlying spectrum, or not being sensitive enough to boost the planet signal of molecules with shallower features. As introduced by Ruffio et al. (2019) we used a likelihood ratio test for the detection of molecules, where we performed a model comparison between an atmospheric model including all species of the input model (Fig. 4) and one including all species except the molecule of interest. Hence we can test if adding a molecule to our atmosphere model significantly fits the data better than omitting it. Calculating a full atmosphere model implies that there are prior constraints for the effective temperature, P–T profile, and gravity of the planet. From self consistent atmospheric models, constraints from existing data and the result of the molecular mapping from Fig. 5 one can obtain an initial estimation for the chemical composition of the atmosphere.

Using this forward model we located the signal in each cube and found its centroid by fitting a 2D Gaussian to the cross-correlation signal. For each sub-band we defined an aperture radius that encompasses 80% encircled energy fraction of the PSF model, and extracted the mean flux within the HPF spaxels of the given aperture centred on the planet location. We then



**Fig. 4.** Continuum subtracted molecule template spectra in the mid-IR at a spectral resolution of the MRS ( $R \sim 3500$  at  $5 \mu\text{m}$ ) that were used in the analysis. The spectra are normalised to the absolute maximum value within the wavelength range of the MRS.

transformed the cross-correlation value into a likelihood function as in Brogi & Line (2019) assuming a normal distribution for the noise:

$$\log(\mathcal{L}) = -\frac{N}{2} \log(\sigma_{\text{data}}^2 - 2R(s) + \sigma_{\text{model}}^2), \quad (2)$$

where  $\sigma_{\text{data}}^2$  and  $\sigma_{\text{model}}^2$  are the variance of the data and the model respectively, and  $R(s)$  is the cross-covariance coefficient for a given wavelength shift  $s$ . The cross-correlation is just the variance normalised cross-covariance. The difference of the log-likelihood of the two models that are being compared yields the log-likelihood ratio (LLR) statistic:

$$\Delta\mathcal{L} = -2(\log(\mathcal{L})_{\text{full}} - \log(\mathcal{L})_{\text{-molecule}}), \quad (3)$$

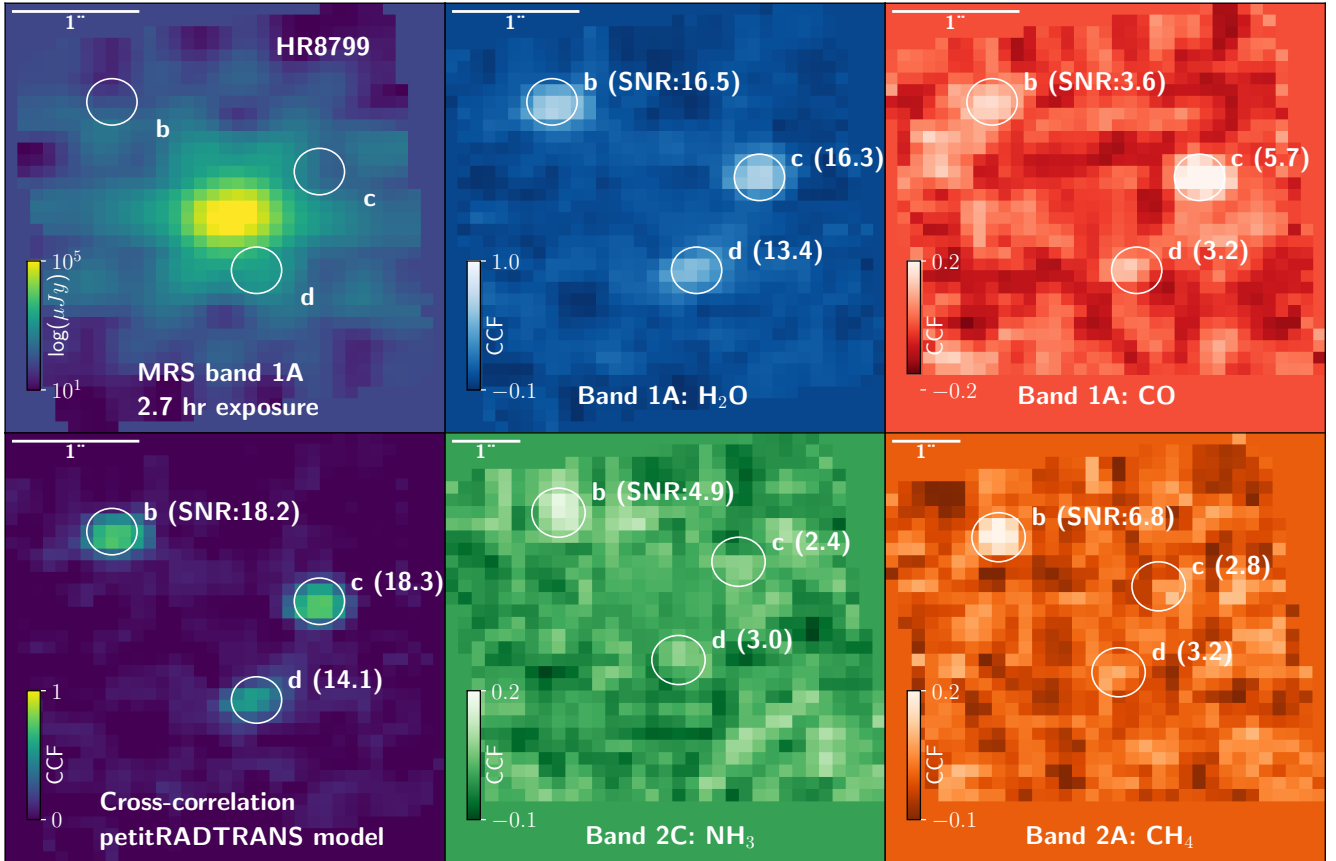
where ‘full’ indicates a forward model with all species and ‘-molecule’ the full model without a given molecule. The resulting test statistic follows a  $\chi^2$  distribution with 1 degree of freedom (d.o.f.), equal to the difference in free parameters of the two models (Brogi & Line 2019).

We proceeded to perform the log-likelihood ratio test in order to test for the existence of molecules in the atmosphere. It is worth noting here that this method will work for the case of having detected the exoplanet in the cube with the forward model using molecular mapping. Moreover, it is only sensitive to atmospheres that contain molecular absorption features to detect since the cross-correlation is not sensitive to continuum modulation. This is especially the case for hotter giant exoplanets, where molecular features become less and less prevalent and are mainly

limited to  $\text{H}_2\text{O}$  and  $\text{CO}$ , or for very cloudy objects, where thick clouds obscure molecular features (Madhusudhan et al. 2016).

For our two simulated data-sets we wanted to inspect which molecules would be detected in the ideal case where we use the input spectrum as the full model. This way we could probe the feasibility of detecting the different molecules, with the only limitation being set by the instrument’s capabilities outlined in Sect. 2.2. Omitting just one molecule and keeping the rest of the molecules constant could still fit better and introduce a bias when applying this to real data, where we do not know the true values of the rest of the parameters. Therefore the LLR statistic values represent an upper limit of the detection sensitivity of the instrument to a specific molecule that is present in the atmosphere.

In Fig. 7 we present the value of the log-likelihood ratio test summing over all sub-bands in Channel 1 and 2 ( $4.8\text{--}12 \mu\text{m}$ ) for a list of molecules that could be expected in such an atmosphere. Values, larger than 25 (critical value such that cumulative distribution of  $\chi^2$  with 1 dof is 0.999997, equivalent to the  $5\sigma$  detection threshold), of the test statistic can be interpreted as molecules with significant contribution to the atmosphere spectrum, conditional on the full forward model that is used as a reference. Values around zero indicate that no significant change in the log-likelihood of the cross-correlation is observed, being the case for molecules that either do not have any features at these wavelengths (e.g.  $\text{CO}$ , only has a feature in sub-band 1A), or weak spectral features that are suppressed due to other more dominant, broadband absorption features present at the same wavelength. This is seen for example of  $\text{CH}_4$  that is almost equally abundant in the two planets but detected at very



**Fig. 5.** Result of molecular mapping for a simulated observation of HR8799 with the MRS for single sub-bands. *First vertical column:* single log-scaled frame from the flux calibrated and reconstructed cube, and the same frame cross-correlated with its input spectrum for sub-band 1A. *Second and third vertical columns:* cross correlation maps of filtered cubes with individual molecular templates correctly identify species of H<sub>2</sub>O, CO for all planets, and CH<sub>4</sub>, NH<sub>3</sub> for planet b. Values in parentheses indicate the S/N and the sub-bands with the highest S/N are plotted.

different strengths due to the presence of much more NH<sub>3</sub> for the colder GJ504 b. Interestingly, PH<sub>3</sub> is detected for HR8799 b despite its low abundance of  $10^{-5}$ . The detection takes place in sub-bands 2B and mainly 2C, and is attributed to the fact that there are many features of PH<sub>3</sub> at these wavelengths that are not suppressed by other molecules with the given composition. This contrasts the non detection of H<sub>2</sub>S which is much more abundant, but its weak spectral signature coincides with CH<sub>4</sub> and NH<sub>3</sub>. Finally, negative numbers of the LLR would indicate that the model without the molecule significantly fits the data better. This is observed in our results in sub-band where multiple dominating absorbers are present simultaneously. The LLR values of each individual sub-band are listed in Tables C.1 and C.2. This approach will work very well with the MRS since it changes the detection of low S/N individual molecular templates to evaluating and comparing models of high S/N of the cross-correlation with the full atmospheric model over the wide wavelength coverage offered by the instrument.

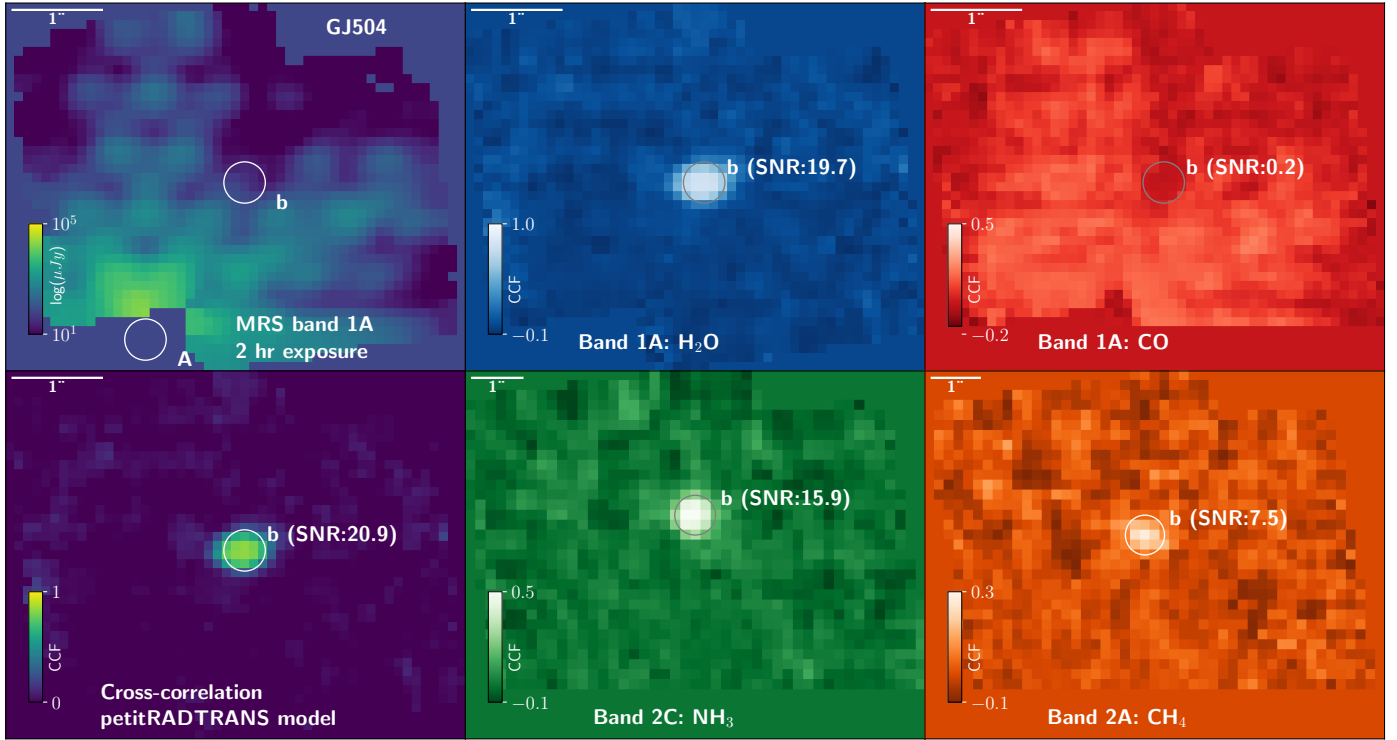
## 5. Discussion

### 5.1. Molecular mapping with the MRS

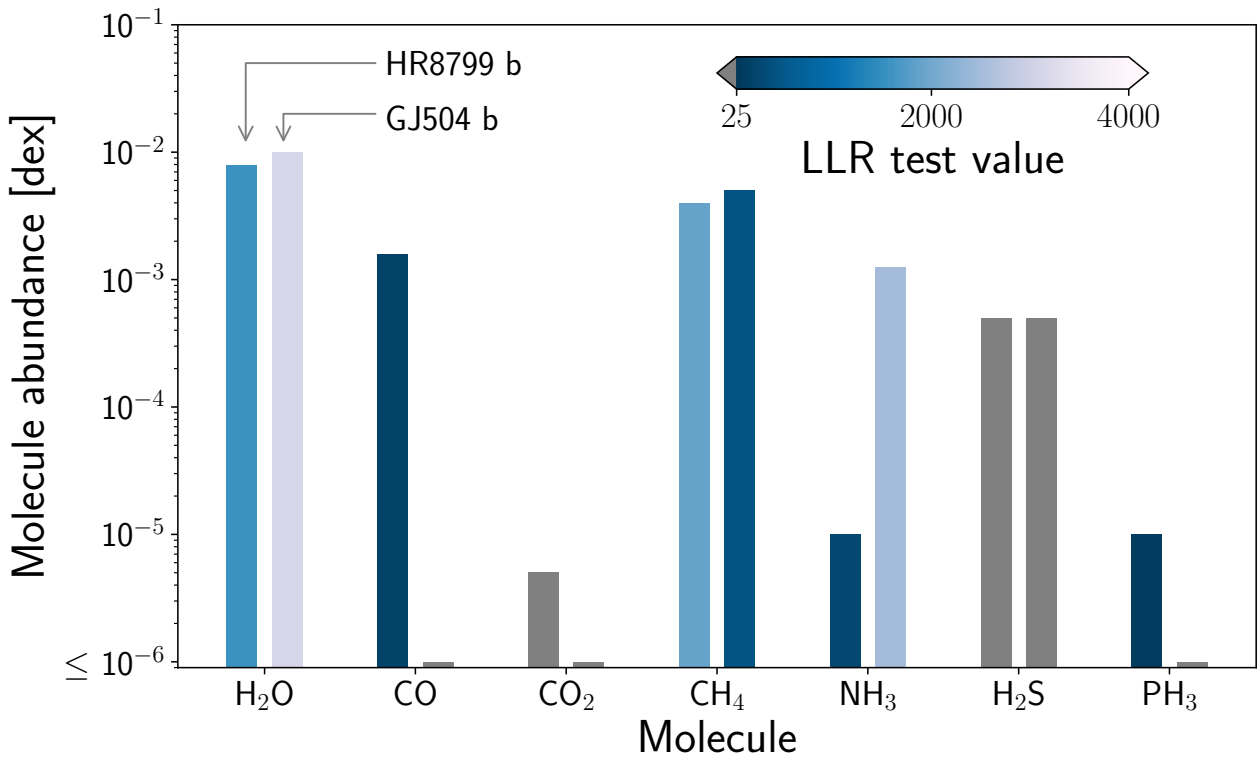
As the only mid-IR IFS until ELT/METIS, the MIRI-MRS is uniquely positioned to contribute to the characterisation of giant exoplanet atmospheres. We show that with the cross-correlation technique, multiple molecular species are detected simultaneously at high confidence for two known systems, taking

advantage of the broad wavelength coverage and moderate spectral resolution. The detection sensitivity of trace species greatly depends on the prior knowledge of the system and will vary as a function of the composition and relative abundances of different species. The mid-IR provides a unique discovery space for constraining the chemical composition of exoplanets. We expect NH<sub>3</sub> being one of the prevalent molecules in the mid-IR, which is supported by our results, being detected in both GJ504 b and HR8799 b. Unexpectedly, PH<sub>3</sub> is also detected in sub-bands 2B and 2C for HR8799 b using the likelihood ratio test, despite its weak molecular signature. Based on chemical equilibrium models, this molecule could be expected in such an atmosphere (Wang et al. 2017) and if proven detectable it would be the first detection of PH<sub>3</sub> in an exoplanet. For planets such as HR8799 b, PH<sub>3</sub> would be well mixed and could trace vertical mixing in the atmosphere (Fletcher et al. 2009), as well as tracing the total phosphorus abundance of the atmosphere (Visscher et al. 2006).

We explored the S/N as a function of the total exposure time of the full model and of different single molecules by using the output of the MRS simulation, only keeping part of the total number of integrations in increasing steps, and re-running the JWST pipeline each time. Because of the dependence of detection on the number of available features and the composition of the atmosphere, the time required to obtain a significant detection with molecular mapping can be very different for each molecule, even for molecules of comparable abundance in the atmosphere. Molecules with many, deep features such as H<sub>2</sub>O and NH<sub>3</sub> are already detected from the first few integrations (in



**Fig. 6.** Result of molecular mapping for a simulated observation of GJ504 b with the MRS for single sub-bands. *First vertical column:* single log-scaled frame from the flux calibrated and reconstructed cube, and the same frame cross-correlated with its input spectrum for sub-band 1A. *Second and third vertical columns:* cross correlation maps of filtered cubes with individual molecular templates correctly identify the abundant species of H<sub>2</sub>O CH<sub>4</sub> and NH<sub>3</sub>. No CO is detected (*top right panel*), as expected from chemical equilibrium models for cold objects as for instance GJ 504 b. Values in parentheses indicate the S/N and the sub-bands with the highest S/N are plotted.



**Fig. 7.** Log-Likelihood Ratio test statistic values for testing the existence of individual molecules in the atmosphere of GJ504 b and HR8799 b. The test statistic is calculated over all sub-bands of Channel 1 and 2. Values larger than 25 indicate a  $5\sigma$  detection for a  $\chi^2$  distribution with 1 d.o.f. The abundances of the individual molecules are taken from the input model shown in Table 2.

a timescale of a few tens of minutes, depending on the contrast of the planet). Molecules such as  $\text{CH}_4$  and  $\text{PH}_3$  can be affected more, both by the abundance of other species and the number of groups/integration possible due to the bright source limit, requiring multiple hours of observations to constrain them well.

Given the simulations presented in this work, we observe that close to the star the stellar photon noise is dominating the cross-correlation signal, while further out the detector noise and photon noise of the given planet are the limiting factors for the cross-correlation S/N. For all separations the spectral resolution of the instrument and the atmospheric composition of the planet in question will greatly impact the S/N, since the cross-correlation signal is dependent on both quantities. To test this we simulated an observation of just the HR8799 planets without the star and measured the S/N. As expected, the S/N was only slightly improved for the outer planet b, while for HR8799 c, d, and e, which are increasingly closer to the star the change in S/N was significant. Removing the star from the simulations allowed for the detection of the HR8799 e planet, that is in reality not considered feasible, and doubled the S/N of the HR8799 d planet. The analysis does not account for systematic errors that the in-orbit state of the MIRI instrument might introduce when observing faint sources or subtle effects the simulations might miss. For example, the inability to obtain deep integrations due to saturation of the detector and a whole series of related effects, such as persistence, that come into play when observing bright stars, might prove a fundamental limit to which targets are observable in an acceptable amount of telescope time.

Regarding the proposed log-likelihood ratio test for the detection of molecular signatures, introducing multiple species complicates the forward modelling and adds a few caveats that should be considered. Assuming a given molecule is actually present in the atmosphere and the full forward model including this molecule produces a high correlation, it can still occur that the model without the molecule produces a higher log-likelihood, hinting towards the molecule not being significant. This could be explained either by the dependence of molecular features on the P-T profile or the presence of another dominating absorber that was not accounted correctly in the initial model. The P-T profile has a large impact on the strength of the molecular feature line-to-continuum ratio and a profile that deviates significantly from the truth could distort the result of the log-likelihood ratio test. In the latter case, if there are two molecules that overlap in a given wavelength range, the removal of one from the model could enhance the features of the other leading to a better fit and an overall higher value of the LLR test statistic. Given these caveats, we should consider the results of the LLR result as an upper limit for the detection of molecules in the atmospheres of the simulated planets.

One of the next frontiers in characterising atmospheres of self-luminous exoplanets in the next years is the modelling of clouds. Whilst clouds have been studied in more detail for brown dwarfs (see references covered in [Mollière et al. 2020](#)), comparing models to exoplanet spectra has been more challenging, but with increasing amounts of effort dedicated towards this problem ([Rajan et al. 2017](#); [Samland et al. 2017](#); [Charnay et al. 2018](#); [Mollière et al. 2020](#); [Burningham et al. 2021](#)). To test how clouds might affect our results we assumed a cloud model as retrieved by [Mollière et al. \(2020\)](#) for the HR8799 e planet, containing  $\text{MgSiO}_3$  and Fe clouds, including scattering. The physics of this model is already included in `petitRADTRANS`, with a few parameters such as the settling parameter  $f_{\text{sed}}$ , the vertical diffusion coefficient  $K_{\text{zz}}$ , cloud base pressures, and cloud mass fraction for each of the cloud species. We see that the given cloud model

dampens the emission spectra of up to 30% in between 2 and 5  $\mu\text{m}$  and for wavelengths higher than 10  $\mu\text{m}$  (clouds affect the spectra below 2  $\mu\text{m}$  a lot more). In the region between 5 and 10  $\mu\text{m}$  where a lot of molecular features are located for the MRS, the spectra seem less affected. When simulating observations of cloudy spectra with MIRISIM, the clouds reduced the S/N of the cross-correlation molecule detection by up to 20%. Molecules such as  $\text{NH}_3$  and  $\text{CH}_4$  are affected more since the clouds show a higher extinction at the wavelengths where these molecules have features.

As is often the case, we can expect that dedicated data-driven algorithms will improve the sensitivity of the instrument to faint sources such as exoplanets (e.g. recent high contrast imaging results with HST by [Zhou et al. 2021](#)). For the MRS there are two avenues of the data reduction that should be explored. First, moving the analysis from the data cube to the detector space can benefit our science case by decreasing interpolation effects and correlated noise introduced by the cube reconstruction algorithm. Second, one of the crucial steps in direct imaging is modelling the stellar PSF and removing it from the data to improve the contrast to fainter companions ([Stolker et al. 2019](#)). This is handled in a very simplistic way in our analysis, as we did not have the ability to simulate realistic behaviour of the instrument PSF. Once in orbit, we will have a better grasp on the effect the optical stability will have on the cross-correlation and can test PSF modelling ideas as for example in [Ruffio et al. \(2019\)](#) or [Ygouf et al. \(2017\)](#).

By design, molecular mapping - especially in space - is quite robust to false positive detection since it relies on very specific patterns in the data that instrument systematic errors cannot replicate easily, contrary to some detector systematic trend in the continuum that could be interpreted as a broad molecule absorption feature. Additionally, compared to ground based observations, contamination due to earth's atmospheric absorption features will not be an issue for the MRS. With upcoming programmes from Guaranteed Time Observations (GTO, [Birkmann et al. 2017](#)) and Early Release Science (ERS, [Hinkley et al. 2017](#)), we will have the opportunity to test this technique on targets with relaxed contrast requirements and compare the results against state of the art Bayesian fitting methods.

## 5.2. Cross-correlation and retrievals

Detecting the existence of molecules in an atmosphere based on the cross-correlation signal is very useful to verify that the planet exhibits the expected trends that are predicted from exoplanet atmosphere theory. However, the end goal is to measure the abundance of a variety of molecules in the atmosphere. With such measurements one can build a deeper understanding of the physical properties of the atmosphere and compute elemental abundance ratios such as C/O and N/O that can both place exoplanets in context with each other and shed light on formation pathways ([Nowak et al. 2020](#)).

Having transformed the cross-correlation into a likelihood function, as in [Brogi & Line \(2019\)](#) we argue that the next logical step is performing a free retrieval within a Bayesian framework. Since it is expected that the high-pass filtering process removes much of the information regarding the temperature and gravity, we would place tight priors on these parameters and let the retrieval focus on estimating the abundances of molecules. This is justified from previous attempts of estimating the temperature and gravity with a grid of models for beta Pictoris b ([Hoeijmakers et al. 2018](#)) and HIP65426 b ([Petrus et al. 2021](#)),

that show that no clear optimum is reached and the cross-correlation tends towards high values for both parameters.

But we are not limited to just using the MIRI-MRS medium resolution spectra to constrain our models. All data, including near-IR spectra, flux measurements, and especially MIRI Coronagraphic measurements, that cover part of the MRS wavelength, can be included as additional likelihood terms (Brogi & Line 2019). With better quality data from instruments such as GRAVITY (GRAVITY Collaboration 2017) and NIRSpec we can expect the retrievals to take advantage of the complementary information that different wavelength ranges and observing modes offer. For example, Burningham et al. 2021 demonstrated how both near and mid infrared data are required to constrain atmosphere and cloud properties.

From Figs. 1 and 4 we can observe that the features of molecules that are probed through molecular mapping are not necessarily coinciding with the broadband molecule features that are currently the focus of atmosphere retrievals. As an example in the mid-IR the  $\text{NH}_3$  feature between 10 and 11  $\mu\text{m}$  is the focus of most observations and hence the MIRI Coronagraphs have filters at 10.65 and 11.4  $\mu\text{m}$  to be able to measure the  $\text{NH}_3$  feature from the colour difference. For molecular mapping  $\text{NH}_3$  has continuous and identifiable features from 5 to 10  $\mu\text{m}$  and the 10–11  $\mu\text{m}$  wavelength range where the broadband absorption feature is located does not show any particularly unique behaviour in the high pass filtered spectra. This is where the complementary information stems from. The retrieval routine can use information from both broadband features and the high frequency features simultaneously to constrain parameters and potentially overcome bias effects, such as using a limited part of the wavelength, that might affect the posterior distribution. The evaluation of the performance and gain in adding mid-IR medium resolution spectra to an atmospheric retrieval analysis will be explored in depth as part of future work.

### 5.3. Sensitivity of cross-correlation to calibration errors

The cross-correlation signal of a planet spectrum with a forward model enables the detection of the planet under the dominant stellar flux. Here, we evaluate the sensitivity of the cross-correlation signal to instrumental effects that could arise with JWST in orbit. While extended MRS calibration campaigns on the ground and during commissioning with JWST/MIRI in space will be able to correct many of the problems that are found in the data, there are uncertainties in the photometry and wavelength calibration that might not be possible to address. For example, fringing of the detectors creates subtle deviations from the calibration model due to the dependence of the fringes to the exact phase of the incoming beam. It has been verified that the uncertainty of these errors remains within acceptable limits, a process which will be repeated when the instrument is in orbit. Furthermore, in case one calibration step, such as fringing or spectro-photometric calibration, in the pipeline is very critical for the proposed science of this work, additional calibration effort might be required. There is already a plan in place to improve the instrument performance during Cycle 1 with dedicated calibration observations<sup>9</sup>.

We model both photometric and wavelength distortion errors with a simple approach, regarding individual, independent spatial frequencies that systematic errors might manifest in, and control the amplitude of the error as a function of the root

mean square error (RMSE) of the residuals. We begin with an exoplanet spectrum model computed with `petitRADTRANS` and convolve it to the mean resolution of each MRS sub-band and bin it to the detector sampling. Then an additive or multiplicative photometric error is applied, or the wavelength coordinates of the spectrum with respect to the forward model is distorted.

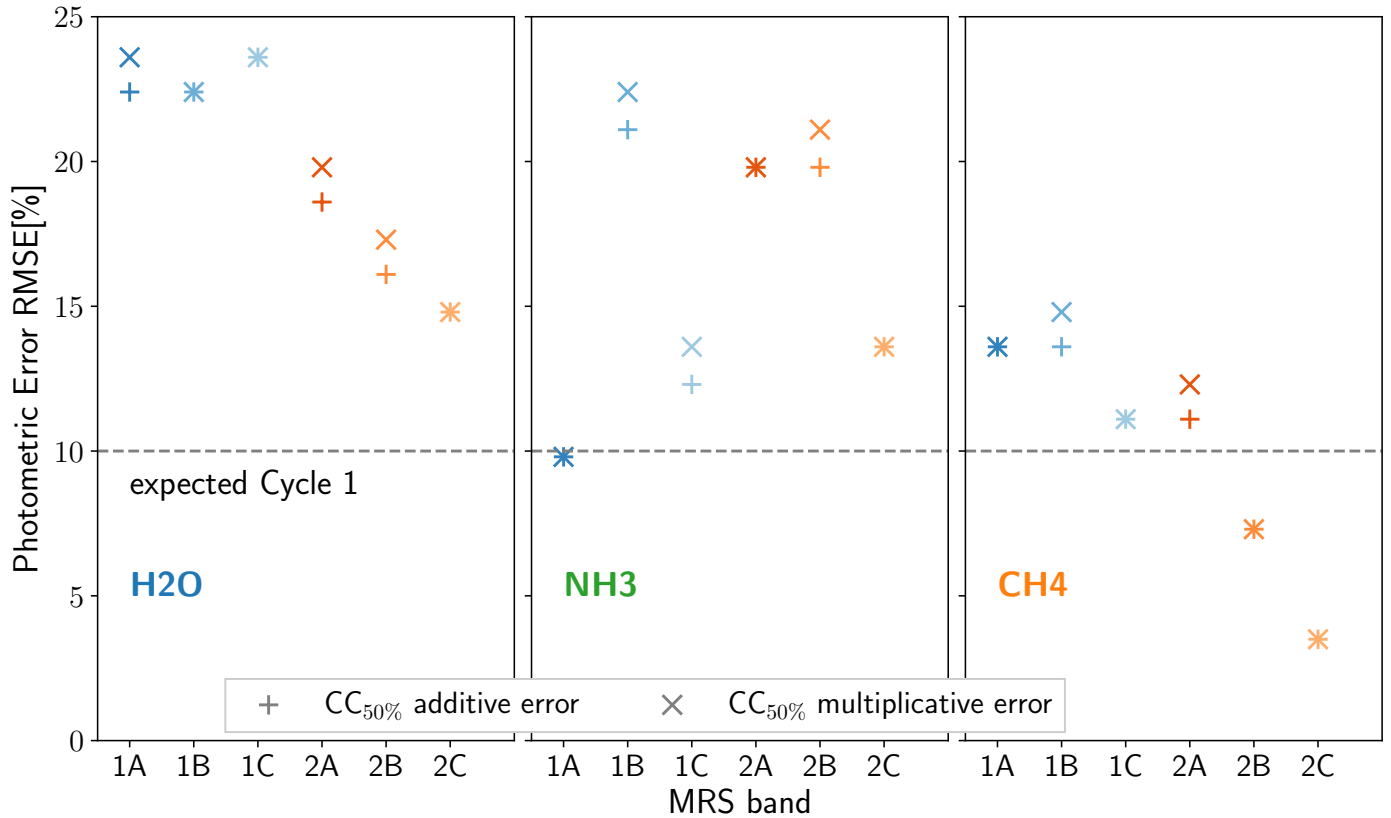
Since the wavelength coordinate is recorded on the detector we consider each wavelength bin to represent a pixel on the detector, and spatial frequencies of detector pixel-wise dependent errors are simply transformed one-to-one onto the wavelength grid of our spectrum. This is not entirely correct since in the MRS detectors wavelength and spatial coordinates are curved on the detector but we consider this approach suitable for tackling this question. We then go through all the spatial frequencies that are constrained within the size of the detector, with the smallest frequency being a gradient over the whole array, and the highest frequency being a modulation with a period of one pixel. In Fig. D.1 examples of the different errors generated are shown.

We then proceed to high pass filter the simulated noisy spectra and cross-correlate them with the noiseless spectrum, measuring the peak value of the cross-correlation. In Fig. 8 we plot the effect of photometric errors on the cross-correlation signal of different molecules for the MRS sub-bands of Channel 1 and 2. For this test we took the single molecule templates seen in Fig. 4 and added the additive and multiplicative photometric noise and calculated the average cross-correlation signal over all the frequencies for each level of noise. We set as a upper limit of photometric error a 50% RMSE, compared to 10% expected in Cycle 1, and plot the RMSE required to decrease the cross-correlation by half. We observe a dependence of the sensitivity for the different molecules as a function the MRS sub-bands, as it is expected since the features change with wavelength. We can conclude that from the tested molecules with the given P-T profile,  $\text{CH}_4$  is the most sensitive to photometric errors. Regarding the wavelength calibration error, we find that uncertainty of the current solution (ground solution  $\sim 0.1$  of resolution element, Labiano et al. 2021) suggests that it does not affect the cross-correlation signal, contributing to maximum a 1–2% reduction. Uncertainty in the absolute wavelength solution again does not affect the result since it would manifest as an error in the radial velocity shift of the planet.

### 5.4. Best targets might yet come

The current limitation of the demographics of directly imaged planets is the small number of detected sub-stellar, planetary-mass companions and two of the published direct imaging surveys (GPI, SHINE Nielsen et al. 2019; Vigan et al. 2021) indicate detection limits as a function of mass and semi-major axis of a few times the mass of Jupiter at tens of AU. Recently, the Young Suns Exoplanet Survey (YSES Bohn et al. 2020; Bohn et al. 2021) discovered new planets at wide orbits as part of a sample of young solar-type stars. Given the large separation from their host, these planets would be prime targets for characterisation with the MRS and more might be discovered by the YSES survey. Carter et al. (2021) investigated the detection limits expected with the JWST coronagraphs and found increased sensitivity to Jupiter and even sub-Jupiter mass, widely orbiting giant exoplanets that are missed due to the limited sensitivity of ground based instruments. If detected, follow up characterisation of these planets could be performed with the MRS, greatly contributing to exoplanet demographics and atmosphere characterisation of a diverse sample of objects.

<sup>9</sup> <https://www.stsci.edu/jwst/observing-programs/calibration-programs>



**Fig. 8.** Sensitivity of cross-correlation to additive and multiplicative photometric errors of single molecule spectra: the plotted points represent the root mean square error with respect to the noiseless spectrum that is required such that the cross-correlation peak is reduced by 50%. The dashed line indicates the expected photometric error in Cycle 1.

## 6. Conclusion

The *James Webb* Space Telescope will open the window to mid-infrared characterisation of giant exoplanet atmospheres at an unprecedented sensitivity. We show that the Medium Resolution Spectrometer on board JWST/MIRI enables the detection of molecules from the emission of directly imaged exoplanets, by taking advantage of the moderate spectral and spatial resolution to disentangle the planet signal from stellar contamination. Constraints derived at these wavelengths can complement other spectroscopic observations in the near-infrared, as well as mid-infrared coronagraphic flux measurements to shed light on the atmosphere physics, chemistry, and formation of these class of planets.

*Acknowledgements.* We thank the anonymous reviewer for their helpful comments that improved the manuscript quality. G.C. and S.P.Q. thank the Swiss National Science Foundation for financial support under grant number 200021\_169131. I.A. thanks the European Space Agency (ESA) and the Belgian Federal Science Policy Office (BELSPO) for their support in the framework of the PRODEX Programme. This research has made use of the NASA Astrophysics Data System and the python packages `numpy` (Harris et al. 2020), `scipy` (Virtanen et al. 2020), `matplotlib` (Hunter 2007) and `astropy` (Astropy Collaboration 2013, 2018).

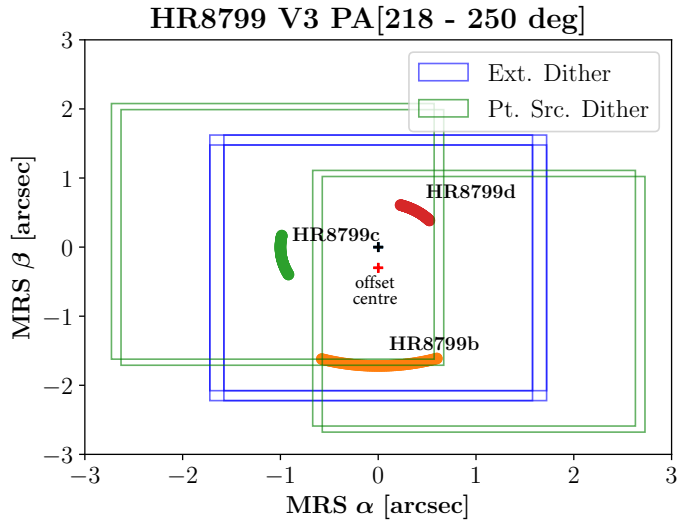
## References

- Allard, F., Hauschildt, P. H., Alexander, D. R., & Starrfield, S. 1997, *ARA&A*, 35, 137
- Argyriou, I., Wells, M., Glasse, A., et al. 2020, *A&A*, 641, A150
- Artigau, É., Gagné, J., Faherty, J., et al. 2015, *ApJ*, 806, 254
- Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, 558, A33
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, 156, 123
- Bacon, R., Accardo, M., Adjali, L., et al. 2010, *SPIE Conf. Ser.*, 7735, 773508
- Bagnasco, G., Kolm, M., Ferruit, P., et al. 2007, *Cryogenic Optical Systems and Instruments XII*, 6692, 66920M
- Barman, T. S., Konopacky, Q. M., Macintosh, B., & Marois, C. 2015, *ApJ*, 804, 61
- Benisty, M., Bae, J., Facchini, S., et al. 2021, *ApJ*, 916, L2
- Beuzit, J. L., Vigan, A., Mouillet, D., et al. 2019, *A&A*, 631, A155
- Birkmann, S., Ferruit, P., & Lagage, P.-O. 2017, *Characterizing the TWA 27 system* (USA: JWST Proposal. Cycle 1)
- Boccaletti, A., Lagage, P. O., Baudino, P., et al. 2015, *PASP*, 127, 633
- Bohn, A. J., Kenworthy, M. A., Ginski, C., et al. 2020, *ApJ*, 898, L16
- Bohn, A. J., Ginski, C., Kenworthy, M. A., et al. 2021, *A&A*, 648, A73
- Bonnefoy, M., Chauvin, G., Rojo, P., et al. 2010, *A&A*, 512, A52
- Bonnefoy, M., Zurlo, A., Baudino, J. L., et al. 2016, *A&A*, 587, A58
- Bonnefoy, M., Perraut, K., Lagrange, A.-M., et al. 2018, *A&A*, 618, A63
- Boss, A. P. 1997, *Science*, 276, 1836
- Bowler, B. P. 2016, *PASP*, 128, 102001
- Bowler, B. P., Liu, M. C., Mawet, D., et al. 2017, *AJ*, 153, 18
- Bozza, V., Mancini, L., & Sozzetti, A. 2018, *Astrophysics of Exoplanetary Atmospheres* (Berlin: Springer), 450
- Brandl, B., Bettonvil, F., van Boekel, R., et al. 2021, *The Messenger*, 182, 22
- Brandt, G. M., Brandt, T. D., Dupuy, T. J., Michalik, D., & Marleau, G.-D. 2021, *ApJ*, 915, L16
- Brogi, M., & Line, M. R. 2019, *AJ*, 157, 114
- Burgasser, A. J., Simcoe, R. A., Bochanski, J. J., et al. 2010, *ApJ*, 725, 1405
- Burningham, B., Faherty, J. K., Gonzales, E. C., et al. 2021, *MNRAS*, 506, 1944
- Carnall, A. C. 2017, *SpectRes: A Fast Spectral Resampling Tool in Python*
- Carter, A. L., Hinkley, S., Bonavita, M., et al. 2021, *MNRAS*, 501, 1999
- Charnay, B., Bézard, B., Baudino, J. L., et al. 2018, *ApJ*, 854, 172
- Cheetham, A. C., Samland, M., Brems, S. S., et al. 2019, *A&A*, 622, A80
- Cugno, G., Patapis, P., Stolker, T., et al. 2021, *A&A*, 653, A12
- Currie, T., Guyon, O., Lozi, J., et al. 2019, *SPIE Conf. Ser.*, 11117, 111170X
- Daemgen, S., Todorov, K., Silva, J., et al. 2017, *A&A*, 601, A65
- Danielski, C., Baudino, J.-L., Lagage, P.-O., et al. 2018, *AJ*, 156, 276
- Davies, R., Esposito, S., Schmid, H. M., et al. 2018, *SPIE Conf. Ser.*, 10702, 1070209

- Delorme, P., Gagné, J., Girard, J. H., et al. 2013, *A&A*, **553**, L5
- De Rosa, R. J., Rameau, J., Patience, J., et al. 2016, *ApJ*, **824**, L21
- Eisenhauer, F., Abuter, R., Bickert, K., et al. 2003, *Proc. SPIE Conf. Ser.*, **4841**, 1548
- Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2016, *A&A*, **595**, A83
- Fletcher, L. N., Orton, G. S., Teanby, N. A., & Irwin, P. G. 2009, *Icarus*, **202**, 543
- Gaia Collaboration (Prusti, T., et al.) 2016, *A&A*, **595**, A1
- Gao, P., Thorngren, D. P., Lee, G. K. H., et al. 2020, *Nat. Astron.*, **4**, 951
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *Space Sci. Rev.*, **123**, 485
- Glasse, A., Rieke, G. H., Bauwens, E., et al. 2015, *PASP*, **127**, 686
- Glauser, A. M., Glasse, A., Morrison, J. E., et al. 2010, *SPIE 7731*, 77313K
- GRAVITY Collaboration (Abuter, R., et al.) 2017, *A&A*, **602**, A94
- Gray, R. O., & Kaye, A. B. 1999, *AJ*, **118**, 2993
- Haffert, S. Y., Bohn, A. J., de Boer, J., et al. 2019, *Nat. Astron.*, **3**, 749
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, **585**, 357
- Hinkley, S., Baraffe, I., Biller, B., et al. 2017, *High Contrast Imaging of Exoplanets and Exoplanetary Systems with JWST* (USA: JWST Proposal ID 1386), Cycle 0 Early Release Science
- Hoeijmakers, H. J., Schwarz, H., Snellen, I. A. G., et al. 2018, *A&A*, **617**, A144
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Janson, M., Brandt, T., Kuzuhara, M., et al. 2013, *ApJ*, **778**, L4
- Kendrew, S., Scheithauer, S., Bouchet, P., et al. 2015, *PASP*, **127**, 623
- Keppler, M., Benisty, M., Müller, A., et al. 2018, *A&A*, **617**, A44
- Klaassen, P. D., Geers, V. C., Beard, S. M., et al. 2021, *MNRAS*, **500**, 2813
- Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, *Science*, **339**, 1398
- Kuzuhara, M., Tamura, M., Kudo, T., et al. 2013, *ApJ*, **774**, 11
- Labiano, A., Azzollini, R., Bailey, J., et al. 2016, *SPIE Conf. Ser.*, **9910**, 99102W
- Labiano, A., Argyriou, I., Alvarez-Marquez, J., et al. 2021, *A&A*, **656**, A57
- Lafrenière, D., Jayawardhana, R., & van Kerkwijk, M. H. 2010, *ApJ*, **719**, 497
- Larkin, J., Barczys, M., Krabbe, A., et al. 2006, *SPIE Conf. Ser.*, **6269**, 62691A
- Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, *AJ*, **154**, 91
- Lightsey, P. A., Knight, J. S., & Golnik, G. 2014, *SPIE Conf. Ser.*, **9143**, 914304
- Liu, D., Blanton, M. R., & Law, D. R. 2020, *AJ*, **159**, 22
- Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, *Proc. Natl. Acad. Sci.*, **111**, 12661
- Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Science*, **350**, 64
- Madhusudhan, N. 2019, *ARA&A*, **57**, 617
- Madhusudhan, N., Agúndez, M., Moses, J. I., & Hu, Y. 2016, *Space Sci. Rev.*, **205**, 285
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Science*, **322**, 1348
- Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010, *Nature*, **468**, 1080
- Mawet, D., Wizinowich, P., Dekany, R., et al. 2016, *Adaptive Opt. Syst.*, **V**, 9909, 99090D
- Miles-Páez, P. A., Metchev, S., Luhman, K. L., Marengo, M., & Hulsebus, A. 2017, *AJ*, **154**, 262
- Mollière, P., Boekel, R. v., Dullemond, C., Henning, T., & Mordasini, C. 2015, *ApJ*, **813**, 47
- Mollière, P., van Boekel, R., Bouwman, J., et al. 2017, *A&A*, **605**, C3
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, **627**, A67
- Mollière, P., Stolker, T., Lacour, S., et al. 2020, *A&A*, **640**, A131
- Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, *ApJ*, **832**, 41
- Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, *AJ*, **158**, 13
- Nowak, M., Lacour, S., Lagrange, A.-M., et al. 2020, *A&A*, **642**, L2
- Öberg, K. I., & Bergin, E. A. 2021, *Phys. Rep.*, **893**, 1
- Öberg, K. I., & Wordsworth, R. 2019, *AJ*, **158**, 194
- Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, *ApJ*, **743**, L16
- Pathak, P., Petit dit de la Roche, D. J. M., Kasper, M., et al. 2021, *A&A*, **652**, A121
- Patience, J., King, R. R., de Rosa, R. J., & Marois, C. 2010, *A&A*, **517**, A76
- Pearce, L. A., Kraus, A. L., Dupuy, T. J., et al. 2019, *AJ*, **157**, 71
- Perryman, M., Hartman, J., Bakos, G. A., & Lindegren, L. 2014, *ApJ*, **797**, 14
- Petit dit de la Roche, D. J. M., Hoeijmakers, H. J., & Snellen, I. A. G. 2018, *A&A*, **616**, A146
- Petit dit de la Roche, D. J. M., van den Ancker, M. E., Kissler-Patig, M., Ivanov, V. D., & Fedele, D. 2019, *MNRAS*, **491**, 1795
- Petit dit de la Roche, D. J. M., Oberg, N., van den Ancker, M. E., et al. 2021, *A&A*, **648**, A92
- Petrus, S., Bonnefoy, M., Chauvin, G., et al. 2021, *A&A*, **648**, A59
- Piso, A.-M. A., Pegues, J., & Öberg, K. I. 2016, *ApJ*, **833**, 203
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *Icarus*, **124**, 62
- Powell, D., Loudon, T., Kreidberg, L., et al. 2019, *ApJ*, **887**, 170
- Rajan, A., Rameau, J., De Rosa, R. J., et al. 2017, *AJ*, **154**, 10
- Rameau, J., Chaussonot, J., Carlotti, A., Bonnefoy, M., & Delorme, P. 2021, *A&A*, **649**, A143
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *J. Astron. Teles. Instrum. Syst.*, **1**, 014003
- Rieke, G. H., Ressler, M. E., Morrison, J. E., et al. 2015a, *PASP*, **127**, 665
- Rieke, G. H., Wright, G. S., Böker, T., et al. 2015b, *PASP*, **127**, 584
- Ruffio, J.-B., Macintosh, B., Konopacky, Q. M., et al. 2019, *AJ*, **158**, 200
- Ruffio, J.-B., Konopacky, Q. M., Barman, T., et al. 2021, *AJ*, **162**, 290
- Samland, M., Mollière, P., Bonnefoy, M., et al. 2017, *A&A*, **603**, A57
- Snellen, I. A. G., & Brown, A. G. A. 2018, *Nat. Astron.*, **2**, 883
- Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, *Nature*, **509**, 63
- Spiegel, D. S., & Burrows, A. 2012, *ApJ*, **745**, 174
- Stolker, T., Bonse, M. J., Quanz, S. P., et al. 2019, *A&A*, **621**, A59
- Stolker, T., Marleau, G. D., Cugno, G., et al. 2020a, *A&A*, **644**, A13
- Stolker, T., Quanz, S. P., Todorov, K. O., et al. 2020b, *A&A*, **635**, A182
- Todorov, K. O., Line, M. R., Pineda, J. E., et al. 2016, *ApJ*, **823**, 14
- Turrini, D., Schisano, E., Fonte, S., et al. 2021, *ApJ*, **909**, 40
- Vigan, A., Fontanive, C., Meyer, M., et al. 2021, *A&A*, **651**, A72
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, **17**, 261
- Visscher, C., Lodders, K., & Bruce Fegley J. 2006, *ApJ*, **648**, 1181
- Wagner, K., Boehle, A., Pathak, P., et al. 2021, *Nat. Commun.*, **12**, 922
- Wakeford, H. R., & Sing, D. K. 2015, *A&A*, **573**, A122
- Wang, D., Miguel, Y., & Lunine, J. 2017, *ApJ*, **850**, 199
- Wang, J. J., Graham, J. R., Dawson, R., et al. 2018, *AJ*, **156**, 192
- Wang, J., Wang, J. J., Ma, B., et al. 2020, *AJ*, **160**, 150
- Wang, J. J., Vigan, A., Lacour, S., et al. 2021, *AJ*, **161**, 148
- Wells, M., Pel, J.-W., Glasse, A., et al. 2015, *PASP*, **127**, 646
- Wilcomb, K. K., Konopacky, Q. M., Barman, T. S., et al. 2020, *AJ*, **160**, 207
- Woitke, P., Min, M., Thi, W. F., et al. 2018, *A&A*, **618**, A57
- Wright, G. S., Wright, D., Goodson, G. B., et al. 2015, *PASP*, **127**, 595
- Wu, Y.-L., Close, L. M., Males, J. R., et al. 2015, *ApJ*, **801**, 4
- Ygouf, M., Beichman, C., Hodapp, K., & Roellig, T. 2017, in *SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, eds. C. Reylé, P. Di Matteo, F. Herpin, E. Lagadec, A. Lançon, Z. Meliani, & F. Royer, Di
- Zhou, Y., Bowler, B. P., Morley, C. V., et al. 2020, *AJ*, **160**, 77
- Zhou, Y., Bowler, B. P., Wagner, K. R., et al. 2021, *AJ*, **161**, 244

## Appendix A: MRS visibility tool

We adapted the coronagraphic visibility tool<sup>10</sup> by rotating the coordinates to match the rotation of the MRS field with respect to the MIRI Imager. Additionally we enable plotting the different available dither patterns and allow for offsets of the target (usually the star). We plan to submit this tool to the developers of the coronagraphic visibility tool.



**Fig. A.1:** Example output of the MRS visibility tool. In order to properly sample the PSF while ensuring all targets in the HR 8799 system remain in view, we choose to use the CH1 extended source (Ext.) 4-point dither pattern, with a  $-0''.3$  offset along the MRS  $\beta$  axis. The nominal point source (Pt. Src.) 4-point dither pattern would provide better spatial sampling but leave one or more planets outside the FoV for half the observations.

<sup>10</sup> <https://jwst-docs.stsci.edu/jwst-other-tools/jwst-target-visibility-tools/jwst-coronagraphic-visibility-tool-help>

## Appendix B: Potential planetary-mass targets for molecular mapping with the MRS

**Table B.1:** Sub-stellar, planetary mass objects that would be observable by MIRI-MRS. The catalogue was retrieved from the Exoplanet.eu ([exoplanet.eu](http://exoplanet.eu)) database and includes all bound objects regarded as planets (excluding verified Brown Dwarfs) with estimated mass up to  $20 M_J$ . Planets that are not resolved from their host star by the MRS and planets orbiting at wide separations ( $>400$  AU) are excluded.

Target	$T_{\text{eff}}$ [K]	a [AU]	Mass [ $M_J$ ]	Distance [pc]	Ang. Sep. [as]	Reference
HR 8799 d	1200	27.0	$8.3 \pm 0.6$	41.3	0.65	Wang et al. (2018)
HR 8799 c	1200	42.9	$8.3 \pm 0.6$	41.3	0.94	Wang et al. (2020)
HR 8799 b	1100	68.2	$7^{+34}_{-2}$	41.3	1.72.	Barman et al. (2015)
GJ 504 b	550	43.5	$4.0^{+4.5}_{-1}$	17.5	2.48	Bonnefoy et al. (2018)
HD 95086 b	1050	61.7	$2.6 \pm 0.4$	90.4	0.62	De Rosa et al. (2016)
HIP 65426 b	1600	109.2	$9 \pm 3$	111.4	0.83	Cheetham et al. (2019)
2M 0219 b	1700	160	$13.5 \pm 1.5$	39.4	3.96	Artigau et al. (2015)
YSES-1 b**	1700	162	$14 \pm 3$	94.6	1.71	Bohn et al. (2020)
YSES-1 c**	1200	320	$6 \pm 1$	94.6	3.4	Bohn et al. (2020)
YSES-2 b	-	115	$6.3^{+1.6}_{-0.9}$	110	1.05	Bohn et al. (2021)
2M 2236 b	1200	230	$12.5 \pm 1.5$	63	3.7	Bowler et al. (2017)
ROXs 42B(AB) b	1900	140	$9 \pm 3$	135	1.2	Daemgen et al. (2017)
HD 203030 b	1050	487	$11^{+4}_{-3}$	41	11.9	Miles-Pérez et al. (2017)
2M 0103(AB) b	-	84	$13 \pm 1$	47.2	1.7	Delorme et al. (2013)
AB Pic b	2000	275	$13.5 \pm 0.5$	47.3	5.5	Bonnefoy et al. (2010)
1RXS 1609 b	1800	330	$14^{+2}_{-3}$	145	2.2	Lafrenière et al. (2010)
CT Cha b**	2500	440	$17 \pm 6$	165	2.6	Wu et al. (2015)
GSC 6214-210 b	1700	320	$17 \pm 3$	145	2.2	Pearce et al. (2019)
51 Eri b*	750	11.1	$2.9 \pm 0.3$	29.4	0.5	Samland et al. (2017)
$\kappa$ And. b*	1850	100	$13^{+12}_{-2}$	50	1.06	Wilcomb et al. (2020)
<b>GTO/ERS Targets</b>						
$\beta$ Pic b	1750	9.2	$9.8 \pm 2.7$	19.3	0.5	Nowak et al. (2020)
2M 1207 b	1600	42	$4^{+6}_{-1}$	52.4	0.78	Patience et al. (2010)
VHS 1256 b	1000	102	$19 \pm 5$	22.2	8.1	Zhou et al. (2020)
Ross 458 c	650	1160	$11.3 \pm 4.5$	63.0	102	Burgasser et al. (2010)

**Notes.** (\*) The host stars of these sub-stellar companions are over the bright source limit of the MRS detectors and can not be observed until partial saturation of the detectors is allowed., (\*\*) JWST GO Cycle 1 accepted proposal to be observed with the MRS or LRS.

## Appendix C: Log-likelihood ratio tables

**Table C.1:** Log-Likelihood Ratio test statistic values for testing the existence of individual molecules in the atmosphere of HR8799 b. Numbers in bold indicate a significant value for a 99.9% ( $3\sigma$ ) confidence level (critical value  $> 10.384$ ).

Molecule	1A	1B	1C	2A	2B	2C	3A	3B	3C	Abundance
H2O	<b>920.6</b>	-192.8	-59.4	-16.3	<b>323.2</b>	<b>22.2</b>	<b>146.5</b>	<b>366.3</b>	<b>39.3</b>	-2.1
CO	<b>184.4</b>	-0.0	0.0	-0.0	-0.0	-0.0	-0.0	-0.0	0.0	-2.8
CO2	-0.2	0.0	0.0	-0.0	0.9	-0.1	2.5	<b>36.5</b>	-21.0	-5.3
CH4	<b>314.1</b>	<b>289.9</b>	<b>585.5</b>	<b>319.3</b>	<b>335.7</b>	<b>57.5</b>	3.3	0.4	-1.9	-2.4
NH3	<b>12.3</b>	<b>22.7</b>	-38.3	-9.9	0.4	<b>139.4</b>	<b>72.4</b>	<b>58.2</b>	4.3	-5.0
H2S	-0.0	0.6	-10.7	9.7	2.3	-0.0	-0.0	-0.0	-1.6	-3.3
PH3	-7.5	0.0	0.3	<b>10.4</b>	<b>44.3</b>	<b>61.1</b>	-0.2	-0.5	-0.0	-5.0

**Table C.2:** Log-Likelihood Ratio test statistic values for testing the existence of individual molecules in the atmosphere of GJ504 b. Numbers in bold indicate a significant value for a 99.9% ( $3\sigma$ ) confidence level (critical value  $> 10.384$ ).

Molecule	1A	1B	1C	2A	2B	2C	Abundance
H2O	<b>1752.2</b>	<b>932.7</b>	<b>321.4</b>	<b>84.4</b>	<b>37.7</b>	-14.9	-2.0
CO	0.0	0.0	0.0	0.0	0.0	0.0	-10.0
CO2	0.0	0.0	0.0	0.0	0.0	0.0	-10.0
CH4	<b>152.0</b>	-65.3	<b>139.6</b>	<b>465.5</b>	-252.5	2.3	-2.3
NH3	<b>245.5</b>	<b>681.0</b>	<b>137.0</b>	<b>339.1</b>	<b>438.2</b>	<b>652.6</b>	-2.9
H2S	0.0	-0.0	<b>15.6</b>	-4.5	-7.9	-0.2	-3.3
PH3	0.0	0.0	0.0	0.0	0.0	0.0	-8.0

## Appendix D: Error modelling

All errors considered here are modelled as sine function of a given frequency  $\nu$  and amplitude  $A$

$$\epsilon(A, \nu) = A \sin(2\pi\nu\lambda)$$

The RMSE for additive errors is related to the amplitude as:

$$RMSE = \frac{|A|}{\sqrt{2}},$$

and is scaled to the maximum of the signal in each sub-band. The maximum is chosen since the data are normalised to the maximum after the high-pass filtering.

Additive photometric errors are considered as modulation in the measured DN/s due to effects independent of the flux on the detector, such as detector odd-even row effect (Rieke et al. 2015a). These errors can simply be added to the spectrum as

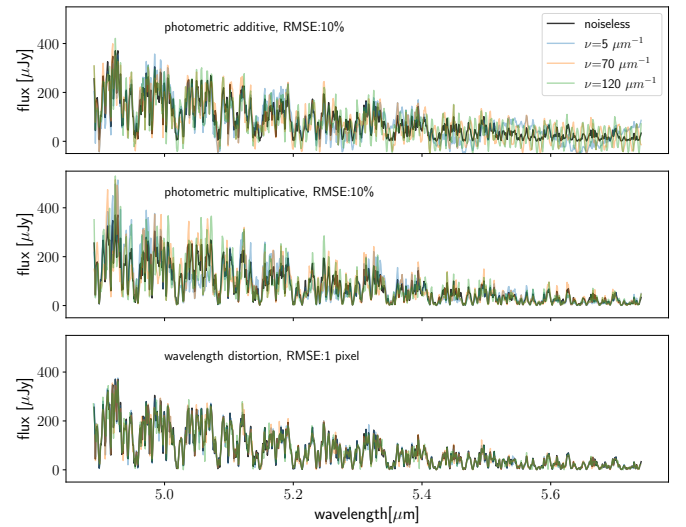
$$f_{additive} = f_{planet} + \epsilon(A, \nu).$$

Multiplicative photometric errors on the other hand have an effect on the transmission of the signal and scale with it. Such effects are similar to detector fringing, detector gain and spectro-photometric conversion uncertainty. These errors are multiplied to the spectrum as a transmission function

$$f_{multiplicative} = f_{planet}(1 + \epsilon(A, \nu)).$$

For these errors, the amplitude was optimised such that the measured RMSE of the residuals was equal to the chosen value. Finally, wavelength distortion errors are again additive systematic errors and they perturb the wavelength coordinates on which the spectrum was sampled. In order to implement this, we took the nominal wavelength grid of each sub-band and added a systematic error creating a new wavelength grid. The high resolution spectrum was then sampled on the new wavelength grid:

$$\lambda' = \lambda + \epsilon(A, \nu).$$



**Fig. D.1:** Example of errors added to the noiseless spectrum with additive (top) multiplicative (middle) photometric errors and wavelength distortion (bottom).