Peering into the formation history of β Pictoris b with VLTI/GRAVITY long-baseline interferometry

Nowak, M.; Lacour, S.; Mollière, P.; Wang, J.; Charnay, B.; Dishoeck, E.F. van; ... ; Ziegler, D.

Peering into the formation history of beta Pictoris b with VLTI/GRAVITY long-baseline interferometry

GRAVITY Collaboration[*]: M. Nowak[1,15], S. Lacour[1,2,3], P. Mollire[9,6], J. Wang[14,**], B. Charnay[1], E.F. van Dishoeck[3,9], R. Abuter[2], A. Amorim[8,11], J.P. Berger[7], H. Beust[7], M. Bonnefoy[7], H. Bonnet[2], W. Brandner[4], A. Buron[3], F. Cantalloube[4], C. Collin[1], F. Chapron[1], Y. Clénet[1], V. Coudé du Foresto[1], P.T. de Zeeuw[3,9], R. Dembet[1], J. Dexter[3], G. Duvert[7], A. Eckart[6,5], F. Eisenhauer[3], N.M. Frster Schreiber[3], P. Fdou[1], R. Garcia Lopez[??,4], F. Gao[3], E. Gendron[1], R. Genzel[3,10], S. Gillessen[3], F. Haumann[3], T. Henning[4], S. Hippler[4], Z. Hubert[1], L. Jocou[7], P. Kervella[1], A.-M. Lagrange[7], V. Lapeyrère[1], J.-B. Le Bouquin[7], P. Léna[1], A.-L. Maire[13,4], T. Ott[3], T. Paumard[1], C. Paladini[2], K. Perraut[7], G. Perrin[1], L. Pueyo[12], O. Pfuhl[3,2], S. Rabien[3], C. Rau[3], G. Rodríguez-Coira[1], G. Rousset[1], S. Scheithauer[4], J. Shangguan[3], O. Straub[1,3], C. Straubmeier[6], E. Sturm[3], L.J. Tacconi[3], F. Vincent[1], F. Widmann[3], E. Wieprecht[3], E. Wiezorrek[3], J. Woillez[2], S. Yazici[3,6], and D. Ziegler[1]

(Affiliations can be found after the references)

December 11, 2019

ABSTRACT

Context. Beta Pictoris is arguably one of the most studied stellar systems outside of our own. Some 30 years of observations have revealed a highly-structured circumstellar disk, with rings, belts, and a giant planet: β Pictoris b. However very little is known about how this system came into being.

Aims. Our objective is to estimate the C/O ratio in the atmosphere of β Pictoris b and obtain an estimate of the dynamical mass of the planet, as well as to refine its orbital parameters using high-precision astrometry.

Methods. We used the GRAVITY instrument with the four 8.2 m telescopes of the Very Large Telescope Interferometer to obtain K-band spectro-interferometric data on β Pic b. We extracted a medium resolution (R=500) K-band spectrum of the planet and a high-precision astrometric position. We estimated the planetary C/O ratio using two different approaches (forward modeling and free retrieval) from two different codes (ExoREM and petitRADTRANS, respectively). Finally, we used a simplified model of two formation scenarios (gravitational collapse and core-accretion) to determine which can best explain the measured C/O ratio.

Results. Our new astrometry disfavors a circular orbit for β Pic b ($e = 0.15^{+0.05}_{-0.04}$). Combined with previous results and with Hipparcos/GAIA measurements, this astrometry points to a planet mass of $M = 12.7 \pm 2.2\,M_{\rm Jup}$. This value is compatible with the mass derived with the free-retrieval code petitRADTRANS using spectral data only. The forward modeling and free-retrieval approches yield very similar results regarding the atmosphere of β Pic b. In particular, the C/O ratios derived with the two codes are identical (0.43 ± 0.05 vs $0.43^{+0.04}_{-0.03}$). We argue that if the stellar C/O in β Pic is Solar, then this combination of a very high mass and a low C/O ratio for the planet suggests a formation through core-accretion, with strong planetesimal enrichment.

Key words. Exoplanets – Instrumentation: interferometers – Techniques: high angular resolution

1. Introduction

The ever-increasing number of exoplanet detections (over 4000, at the time of this writing[1]) proves that our instrumental capabilities are getting better and better at discovering these other worlds. But even though exoplanets are now routinely being observed, determining their physical properties (temperature, mass, composition), let alone the history of their formation, remains extremely challenging. And yet, these measurements are key to understanding the details of planetary formation processes.

Among all measurable quantities, element abundance ratios are emerging as some of the most promising for understanding planetary formation. The question of the supersolar abundances of heavy elements in the atmosphere of Jupiter is probably what motivated the first attempts to link abundance ratios to planetary formation, and several studies have been carried out to understand how planetesimal accretion can lead to heavy element enrichment (Helled et al. 2006; Helled & Schubert 2009; Owen et al. 1999; Alibert et al. 2005). On the exoplanet front, the work of Öberg et al. (2011) was the first general attempt to show that element ratios in an exoplanet atmosphere can be an imprint of its formation history. This idea has since been investigated further by several authors (e.g., Ali-Dib et al. 2014; Thiabaud et al. 2014; Helling et al. 2014; Marboeuf et al. 2014a,b; Madhusudhan et al. 2014, 2017; Mordasini et al. 2016; Öberg & Bergin 2016; Cridland et al. 2016; Eistrup et al. 2016, 2018). While Öberg et al. (2011) highlighted how gas disk abundances can influence the atmospheric composition, the importance of icy planetesimals for the atmospheric enrichment is stressed in Mordasini et al.

(2016), where exoplanet spectra are derived from modeling full formation in the core-accretion paradigm.

Measuring the element ratios is not easy, and requires high-quality data. Madhusudhan et al. (2011) used a free retrieval method on a set of Spitzer and ground-based photometric data in 7 different bands to obtain the first exoplanetary C/O ratio on the hot Jupiter WASP-12b. But the value of C/O > 1 they obtained has since been ruled out by Kreidberg et al. (2015), showing the difficulty of obtaining reliable abundance ratios. Konopacky et al. (2013) used a different approach in their study of HR 8799 c. They obtained K-band spectroscopic observations of the planet with the spectrograph OSIRIS on the Keck II telescope, and were able to extract an estimate of the C/O ratio using model grid fitting. They found a value of C/O = 0.65±0.15. Looking at the same planetary system, Lavie et al. (2017) estimated the C/O ratio for four planets (HR 8799 b, c, d, and e), using a retrieval analysis method. In their analysis, they notably emphasized the importance of high-quality K-band spectroscopic data, which they found to be critical for a reliable measurement of the C/O and C/H ratios.

With the recent direct detection of the giant planet HR 8799 e with the GRAVITY instrument (Gravity Collaboration et al. 2019) on the Very Large Telescope Interferometer (VLTI), optical interferometry has become a new arrow in the quiver of exoplanet observers. By taking advantage of the angular-resolution offered by 100+ meter baselines, optical interferometers can separate a dim exoplanet from the overwhelming residual starlight, leading to accurate measurements of the astrometric position (up to $10\,\mu$as, Gravity Collaboration et al. 2018), and high signal-to-noise spectroscopic data with absolute calibration of the continuum.

In this paper, we present observations of the giant planet β Pic b obtained with GRAVITY and we investigate the possibility of using this K-band spectro-interferometric data to determine the C/O ratio of the planet. The observations are presented in Section 2, together with a brief summary of the data reduction (a complete explanation is given in Appendix A). Section 3 focuses on the orbit and mass of β Pic b. We show in this section how the new GRAVITY astrometric data impacts the best orbital estimate currently available and we provide a new estimate of the dynamical mass of the planet. Section 4 is devoted to the measurement of atmospheric properties and, in particular, to the determination of the C/O ratio, using two different approaches: forward modeling and free retrieval. In Section 5, we discuss the C/O ratio obtained in the case of a formation of β Pic b through gravitational accretion and then through core-accretion. Our general conclusions can be found in Section 6.

2. Observations and data reduction

2.1. Observations

Observations of β Pictoris b were obtained on September, 22, 2018, using the GRAVITY instrument (Gravity Collaboration et al. 2017), with the four 8 m Unit Telescopes (UTs) of the VLT. The instrument was set up in its medium resolution mode (R = 500), and observations were conducted in on-axis/dual-field mode.

The observing strategy was similar to the one described in Gravity Collaboration et al. (2019): the fringe-tracker (Lacour et al. 2019) was using the flux from the central star during the observing sequence, while the position of the science fiber was changed at each exposure, alternating between the central star and the position of the planet. Since the planet was not visible on the acquisition camera, the position used to center the fiber during the planet exposures was a theoretical position, based on predictions from previous monitoring (Wang et al. 2016; Lagrange et al. 2018).

A total of 16 exposures (resp. 17) were acquired on the star (resp. the planet). Each star exposure was made of 50 individual 0.3 s integrations. For the planet, which is ~ 10 mag fainter than the star, the integration time was initially set to 30 s, with 10 integrations per exposure, and reduced to 10 s with 30 integrations at mid-course, since the observing conditions were good (seeing $< 0.8''$). The complete dataset contains 1.4 hr of integration on the planet (and 0.35 hr of associated background exposures), and 4 min 30 s of integration on the central star (plus 1 min 15 s of sky background). The observing log is given in Table 1.

2.2. General data reduction

During planet exposures, the science fiber at each telescope is kept at an offset position with respect to the star, reducing significantly the star to fiber coupling ratio. But even though most of the stellar flux is rejected, speckle noise can still couple to the science fiber and dominate the exposures, hence the need for careful data reduction to disentangle the planet signal from the remaining coherent stellar flux.

The general data reduction method used to reduce the VLTI/GRAVITY observations of β Pic b is presented in details in Appendix A. It can be divided into different parts: pipeline reduction (common to all GRAVITY observations), astrometric extraction, and spectrum extraction. These steps are described in Appendix A.2, A.4, and A.5. The end products are an astrometric position for the planet with respect to the star $(\Delta\alpha, \Delta\delta)$, and a planet-to-star contrast spectrum $C(\lambda) = S_{\mathrm{P}}(\lambda)/S_{\star}(\lambda)$ which is the ratio between the spectra of the planet and of the star.

2.3. K-band spectrum

The contrast spectrum of β Pic b was converted to an absolute spectrum of the planet using a model of the stellar spectrum: $S_{\mathrm{P}}(\lambda) = C(\lambda) \times S_{\star}(\lambda)$. We used a BT-NextGen model (Hauschildt et al. 1999), with a temperature of 8000 K, a surface gravity of $\log(g/g_0) = 4$, and a Solar metallicity, as close as possible to the measured value for this star (Lanz et al. 1995; Gray et al. 2006). We scaled this synthetic spectrum to an ESO K-band magnitude of 3.495, taking into account the correct filter (van der Bliek et al. 1996). This strategy, based on the extraction of a contrast spectrum and the use of a model for the star, helps to reduce the impact of Earth's atmosphere on the final planet spectrum. The result is given in Figure 1.

2.4. Astrometry

Using the data reduction method described in Appendix A.4, we found a mean relative planet to

Target	Start Time (UTC)	End Time (UTC)	EXP	DIT (s)	NDIT	Seeing (")	τ_0 (ms)	Airmass	Parallactic angle (deg)
β Pictoris b	07:37:40	08:31:40	7	30.0	10	0.4/0.9	4.7 / 10.4	1.33 / 1.21	-66.4 / -50.1
SKY	07:50:30	08:24:56	2	30.0	10	0.4 / 0.9	4.7 / 10.4	1.33 / 1.21	N/A
β Pictoris b	08:38:31	09:51:49	10	10.0	30	0.6 / 1.2	5.9 / 8.4	1.20 / 1.12	-47.7 / -16.6
SKY	08:50:41	09:25:03	2	10.0	30	0.6 / 1.2	5.9 / 8.4	1.20 / 1.12	N/A
β Pictoris A	07:43:55	09:58:31	18	0.3	50	0.4 / 1.2	4.7 / 10.4	1.31 / 1.12	-64.7 / -13.2
SKY	07:57:14	09:59:20	5	0.3	50	0.4 / 1.2	4.7 / 10.4	1.31 / 1.12	N/A

Table 1. Observing log for the DDT β Pic b program, carried out on September, 22, 2018.
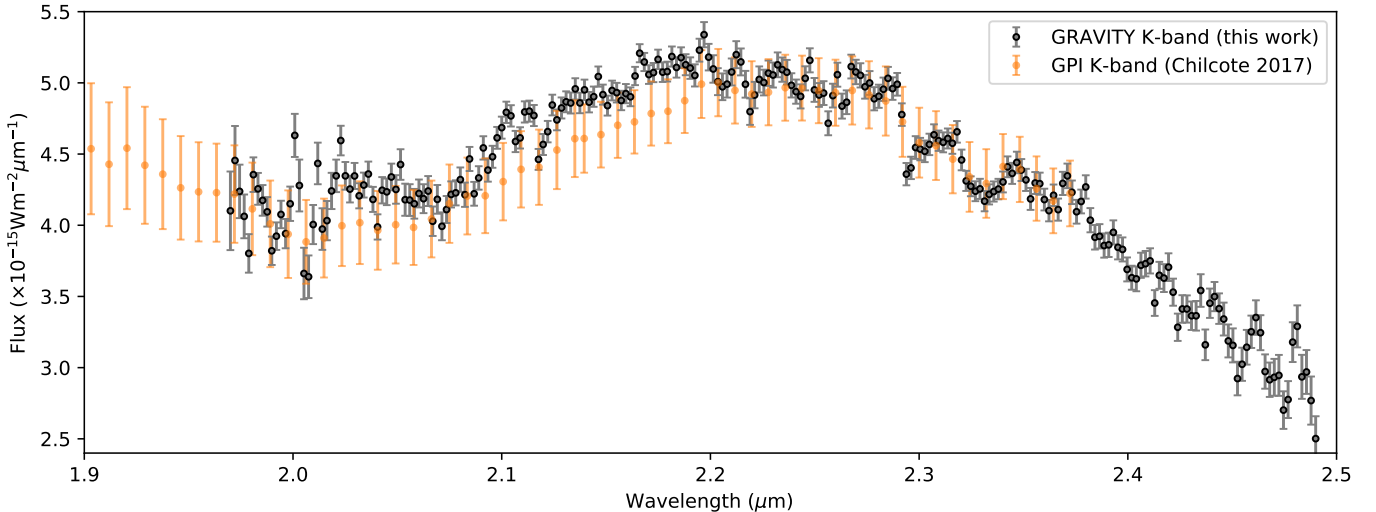

Fig. 1. Calibrated K-band spectrum of β Pictoris b, at R = 500, extracted from the VLTI/GRAVITY observations (gray points). For comparison, the K-band part of the GPI spectrum from Chilcote et al. (2017) (R $\simeq$ 70) is also overplotted (orange points). The error bars plotted for the GRAVITY spectrum only represent the diagonal part of the full covariance matrix.

star astrometry on all the exposure files of:

$$\begin{cases} \Delta\mathrm{RA} = 68.48 \text{ mas} \\ \Delta\mathrm{DEC} = 126.31 \text{ mas} \end{cases} \quad (1)$$

The 1σ confidence interval is given by the covariance matrix of all the 17 exposure files:

$$\mathrm{Covar}\,(\Delta\mathrm{RA}, \Delta\mathrm{DEC}) = \begin{bmatrix} 0.0027 & -0.0035 \\ -0.0035 & 0.0045 \end{bmatrix} \text{ mas}^2$$

This GRAVITY measurement is shown in the inset plot of Figure 2.

In its dual-field mode, GRAVITY is limited to observations of planets above the diffraction limit of a single telescope (to separate the planet from the central star), but the relative astrometry derived from these observations still fully benefits from the length of the telescope array.

3. Orbit and dynamical mass

3.1. Orbital parameters

We fit a Keplerian orbit to the visual astrometry of the planet to characterize its dynamics. As our new GRAVITY point is more than an order of magnitude more precise than any other published astrometric point on the northeastern half of its orbit (c.f., Lagrange et al. 2019a), we expected

a better constraint on the eccentricity of the planet's orbit. We used the published astrometry from Chauvin et al. (2012), Nielsen et al. (2014), and Wang et al. (2016) in this analysis. The orbit was fit using the open-source Python orbit fitting package `orbitize!` (Blunt et al. 2019). We included a custom likelihood to fit the GRAVITY measurement along the two principal axes of the error ellipse. We fit for the same eight parameters as Wang et al. (2016): semi-major axis (a), eccentricity (e), inclination (i), argument of periastron (ω), position angle of the ascending node (Ω), the first periastron passage after MJD = 55,000 in units of fractional orbital period (τ), system parallax, and total system mass (M_{tot}). We generally used relatively unconstrained priors for most of the orbital parameters (see Table 2). For Ω, we constrained it to between $\pi/10$ and $\pi/2$ to account for the fact that Snellen et al. (2014) detected the RV signal of the planet. However, we chose not to explicitly include the RV in the fit as there could be systematics in the reported uncertainties. For the parallax, we used a normal distribution to represent the parallax of 51.44 ± 0.12 mas measured by Hipparcos (van Leeuwen 2007). We sampled the posterior using the parallel-temperature affine-invariant sampler in `ptemcee` (Foreman-Mackey et al. 2013; Vousden et al. 2016) with 20 temperatures, 1000 walkers per temperature. We discarded the first 15,000 steps to allow the walkers to converge. We assessed convergence using the autocorrelation time and by visual inspection of the samples. We then ran each walker for 5000 steps, keeping only ev-
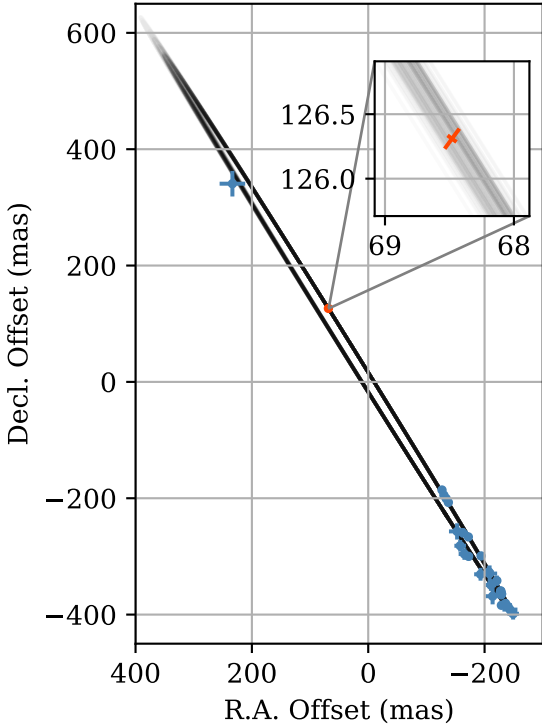
Fig. 2. Visual orbit of β Pic b. Plotted in black are possible orbits randomly drawn from the posterior using only relative astrometry (Section 3.1). Previous astrometric measurements used in the orbit fit are in blue. The GRAVITY measurement from this work is in red, with an inset plot that is zoomed in by a factor of ~ 2000 to display the uncertainties on this measurement.

ery tenth sample to mitigate correlations in the samples produced by any given walker.

Our constraints on the orbit of β Pic b using just astrometry of the planet are collected in Table 2 and plotted in Figure 2. We find that $< 2\%$ of allowed orbits have $e < 0.05$ and $< 0.5\%$ of orbits have $e < 0.03$, although there are still some allowed circular orbits. Dupuy et al. (2019) also proposed an $e \approx 0.25$ when including astrometric and radial velocity data on the system. To statistically assess whether eccentric orbits are preferred, we refit the orbit fixing $e = 0$ and $\omega = 0$ resulting in a fit with two less parameters. Similar to (Wang et al. 2018) in assessing the coplanarity of the HR 8799 planets, we compared the Bayesian Information Criterion (BIC) of the fit that allowed eccentric orbits with the fit that fixed the orbit to be circular, and found that the BIC disfavors the circular orbit by 9.9. The reduction in model parameters for a purely circular orbit does not compensate for an increase in fitting residuals, so we disfavor circular orbits for a single planet model. However, additional confusion on this measurement could be due to a second planet in the system (Lagrange et al. 2019b). The second planet β Pic c would induce epicycles in the apparent orbit of β Pic b around the star due to the gravitational influence of the second planet on the orbit of the host star. Using parameters for β Pic c from

Lagrange et al. (2019b), the magnitude of these epicycles are several hundred μas, so well detectable by GRAVITY, but hidden beneath the uncertainty of previous astrometry. Thus, they would also bias this single GRAVITY measurement, and continued astrometric monitoring is required to separate out the signal of the separate planet from a possibly eccentric orbit of β Pic b.

However, a moderate eccentricity would fit nicely in the dynamics of the system. An $e \approx 0.15$ is consistent with the picture of an eccentric β Pic b launching small bodies towards the star, causing spectroscopic and transiting signatures of exo-comets in observations of the star (Thébault & Beust 2001; Zieba et al. 2019). An interesting question is how such a massive planet acquired a significant eccentricity. The obvious conclusion would point to a second massive planet in the system, such as the radial velocity detected β Pic c (Lagrange et al. 2019b). Otherwise, Dupuy et al. (2019) proposed that if the planet had formed further out and migrated inwards, resonant interactions with the circumstellar disk could pump up its eccentricity to the values we observe today. Characterizing the detailed structure of the circumstellar dust in the system as well as the chemical composition of β Pic b could test this theory.

Generally, the other orbital parameters of β Pic b have already been sufficiently well constrained previously that out results agree with the conclusions drawn in previous works (Millar-Blanchaer et al. 2015; Wang et al. 2016; Lagrange et al. 2019a; Dupuy et al. 2019). We still find that the planet did not transit the star in 2017, and that the Hill sphere of the planet did transit. Assuming a planet mass of 12.9 ± 0.2 $M_{\rm Jup}$, we find a Hill sphere ingress at MJD 57852 ± 2 (2017 April 8) and a Hill sphere egress at MJD 58163 ± 2 (2018 February 13). The closest approach, which does not require an assumption on the planet's mass, is at MJD 58008 ± 1 (2017 September 11), with the planet passing 8.57 ± 0.13 mas from the star (0.166 ± 0.003 au in projection). The precise astrometry of the GRAVITY epoch post conjunction has significantly improved the transit ephemeris from Wang et al. (2016).

3.2. Dynamical mass determination

A significant astrometric acceleration for the star β Pic was detected when comparing its average velocity over the course of the Hipparcos mission and the average velocity inferred by the change in position of the star between the Hipparcos and Gaia missions (Snellen & Brown 2018; Kervella et al. 2019). Assuming this acceleration is due entirely to β Pic b, Snellen & Brown (2018) and Dupuy et al. (2019) used it in conjunction with the visual orbit to measure a dynamical mass for the planet. Snellen & Brown (2018) used the Hipparcos intermediate astrometric data (IAD; van Leeuwen 2007) and Gaia DR2 position (Gaia Collaboration et al. 2018) to fit the position, proper motion, and orbital motion of the host star to derive the mass of the planet. Dupuy et al. (2019) used the re-calibrated Hipparcos and Gaia proper motions from the Hipparcos-Gaia Catalog of Accelerations (HGCA; Brandt 2018) and the stellar radial velocities from Lagrange et al. (2012) to derive the mass of the planet. Being agnostic to which method is more accurate, we repeated both analyses here, now with the new GRAVITY epoch providing strong constraints on a and e, which are otherwise degenerate with the mass of β Pic b, M_b. To repeat the Snellen & Brown

Orbital Element	Prior	Only Relative Astrometry		Hipparcos IAD and Gaia DR2		Brandt (2018) HGCA and Stellar RVs	
		68% CI	Best Fit	68% CI	Best Fit	68% CI	Best Fit
a (au)	LogUniform(1, 100)	10.6 ± 0.5	10.9	$11.0^{+0.3}_{-0.4}$	11.2	$10.0^{+0.6}_{-0.5}$	10.2
e	Uniform(0, 1)	$0.15^{+0.04}_{-0.05}$	0.18	$0.19^{+0.02}_{-0.03}$	0.21	0.11 ± 0.05	0.13
i (°)	$\sin(i)$	89.04 ± 0.03	89.05	89.06 ± 0.02	89.07	$88.99^{+0.03}_{-0.04}$	89.00
ω (°)	Uniform(0, 2π)	196^{+3}_{-4}	196	197 ± 2	197	202 ± 5	202
Ω (°)	Uniform($\pi/10$, $\pi/2$)	31.88 ± 0.05	31.90	31.90 ± 0.05	31.92	31.87 ± 0.05	31.88
τ	Uniform(0, 1)	0.159 ± 0.009	0.157	$0.155^{+0.008}_{-0.006}$	0.152	$0.185^{+0.019}_{-0.016}$	0.185
Parallax (mas)	$\mathcal{N}(51.44, 0.12)$	51.44 ± 0.12	51.45	51.44 ± 0.12	51.49	51.44 ± 0.12	51.47
M_{tot} ($M_\odot$)	Uniform(1.4, 2)	1.82 ± 0.03	1.82	1.83 ± 0.03	1.81	1.79 ± 0.03	1.78
M_b ($M_{\rm Jup}$)	Uniform(1, 100)	-	-	12.7 ± 2.2	13.8	$14.2^{+3.7}_{-3.9}$	15.1

Table 2. Orbital Parameters of β Pic b. Listed are fits using just astrometry of the planet (Section 3.1) and also including measurements of the stellar orbit for dynamical mass estimates of the planet (Section 3.2). For each fit, the first column lists the 68% credible interval centered about the median. The second column lists the fit with the maximum posterior probability. We note that this the best fit orbit is generally not the best estimate of the true orbit. However, it is useful as a valid representative orbit, whereas using the median of all of the orbital parameters often is not a valid orbit due to complex covariances.
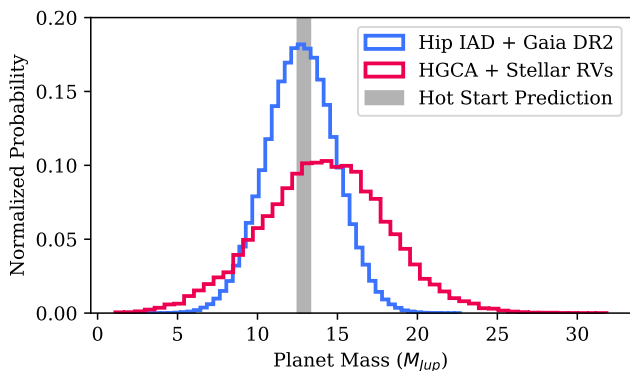

Fig. 3. Dynamical mass estimates of β Pic b using the two different methods described in Section 3.2. The shaded grey region is the 2σ uncertainty on the hot-start derived mass from Chilcote et al. (2017).

(2018) orbit fit, we include five additional parameters in the fit: the position and proper motion of the star in RA and DEC, as well as the mass of the planet. We also switch the prior on parallax to a uniform prior between 50.24 and 52.64 mas, since the Hipparcos intermediate astrometric data now constrains this parallax. To repeat the Dupuy et al. (2019) analysis, we only fit for changes in the tangential velocity of the host star, so we do not need to fit for its actual position and proper motion. We only include a RV offset and RV jitter term for the stellar RV data. We modified the `orbitize!` custom likelihood function to include these measurements of the host star, and repeated the orbit fit.

We list the orbital and mass constraints in Table 2, marginalizing over astrometric parameters of the host star and stellar RV calibration numbers in the two fits. We also plot the posterior probabilties for the mass of β Pic b in Figure 3. In the fit using the Hipparcos IAD, the semi-major axis and eccentricity posteriors now favor slightly higher values by 1σ. We find a dynamical mass of β Pic b of 12.7 ± 2.2 $M_{\rm Jup}$, which is consistent with the values from Snellen & Brown (2018). Conversely, using the recalibrated

stellar astrometry from the Brandt (2018) HGCA catalog and the stellar RVs, we find a slightly lower semi-major axis and eccentricity by 1σ than the relative astrometry only fit. Despite these minor differences, all three fits considered in this work favor an eccentricity between 0.1-0.2. We do not find orbital solutions with $e > 0.25$ as has been suggested by Dupuy et al. (2019). In the HGCA fit, we also find a weaker dynamical mass constraint for β Pic b of $14.2^{+3.7}_{-3.9}$ $M_{\rm Jup}$, which is consistent with Dupuy et al. (2019). The Hipparcos IAD method provides more stringent constraints on the planet mass, likely because it has smaller uncertainties. It is unclear whether this better constrain is unbiased, or if the uncertainties are underestimated due to calibration systematics or effects of other planets on the stellar astrometry. However, as seen in Figure 3, both fits agree with each other, and both dynamical masses are consistent with hot-start derived masses of 12.7 ± 0.3 $M_{\rm Jup}$ from Morzinski et al. (2015) and 12.9 ± 0.2 $M_{\rm Jup}$ from Chilcote et al. (2017). More accurate stellar astrometry or RVs are necessary to test hot-start evolutionary models more stringently, given that the model-dependent hot-start masses have an order of magnitude better precision than the dynamical masses.

4. The atmosphere of beta Pic b

4.1. Previous work

Physical parameters of β Pictoris b have been reported in a number of previous studies (see Table 15 of Morzinski et al. (2015), Table 2 in Chilcote et al. (2017) for a summary of these results). The temperature of the planet has been estimated by several authors, using atmospheric or evolutionary model grid fitting, on photometric and/or spectroscopic data. The most extensive study to date was performed by Chilcote et al. (2017), who obtained GPI spectroscopic data at $R \simeq 50$ in Y, J, H, and K-band, as well as photometric points in different bands ranging from 1 to 5 μm. Using different atmospheric models (BT-Settl: Allard et al., 2012; DRIFT-PHOENIX: Woitke & Helling, 2003, Helling & Woitke, 2006; AMES-DUSTY: Chabrier et al., 2000; Allard et al., 2001), they obtained values ranging from 1650 K to 1800 K for the temperature, and 3.0 to 4.5 for

$\log(g/g_0)$. These values are similar to what is reported in Bonnefoy et al. (2013, 2014), Chilcote et al. (2015), and Morzinski et al. (2015), with the same models.

The lower limit of the range of temperatures estimated comes from Baudino et al. (2015). Using their Exo-REM model grid, and a set of photometric data only (the GPI spectrum was not available at the time), they derived a temperature of 1550 K, and a surface gravity of $\log(g/g_0) = 3.5$.

4.2. ExoREM atmospheric grid fitting

Using either the GRAVITY K-band spectrum only, or the GRAVITY K-band and GPI YJH bands spectra, we performed a grid model fitting using the newest ExoREM grid (Charnay et al. 2018).

We performed a χ^2-based grid model fitting on the GRAVITY K-band only data, using the same ExoREM model grid as used to fit the GRAVITY HR 8799 e spectrum in Gravity Collaboration et al. (2019), ranging from 400 to 1800 K in temperature, with a step-size of 50 K, from 3.0 to 5.0 in $\log(g/g_0)$, with a step-size of 0.2, for a metalicity of [Fe/H] = -0.5, 0, and 0.5, and with a Solar C/O ratio. The best fit was obtained for a Solar metallicity, a temperature of 1750 K, and a $\log(g/g_0)$ of 3.30. However, this best fit also leads to a mass of $1.3\,M_{\rm Jup}$, more than $5\,\sigma$ away from our estimate given in Table 2. To force the result of the fit to be in agreement with our mass estimate, we added a mass prior in the χ^2 calculation. We used a weight for the prior similar to the weight of the entire GRAVITY spectrum:

$$\chi^2 = n_\lambda \frac{(m - 12.7\,M_{\rm Jup})^2}{(2.2\,M_{\rm Jup})^2} + \sum \frac{(F_{\rm data}(\lambda_k) - F_{\rm model}(\lambda_k))^2}{\sigma_F(\lambda_k)^2} \quad (2)$$

in which $F_{\rm data}$ and $F_{\rm model}$ represent the flux from the data and from the model at the different wavelengths, σ_F the error on the data, and m the mass derived from the flux level.

With this new definition of the χ^2, the same ExoREM grid led to a best fit at $T = 1500$ K, $\log(g/g_0) = 4.0$, for a Solar metallicity. The corresponding planet radius is $1.9\,R_{\rm Jup}$, and the mass is $14\,M_{\rm Jup}$, compatible with the estimate of Section 3. However, we find that the fit itself was not very good, with a $\chi^2_{\rm red}$ value of 6.8. The CO region around 2.3 μm was particularly poorly fitted.

The fit was improved by generating a second ExoREM grid, which included the C/O ratio as an additional parameter. The grid was generated on the same range of temperature, surface gravity, and metallicity, for C/O values ranging from 0.3 to 0.8, with a step of 0.05.

Without the mass prior, the new grid yielded a best fit corresponding to a temperature of 1700 ± 50 K, a surface gravity of $\log(g/g_0) = 3.5$, a metallicity of -0.5 (the lowest value available in our grid), and a C/O ratio < 0.3. The resulting planet mass remained too low, at $2\,M_{\rm Jup}$. Adding the mass prior in the definition of the χ^2, as in Equation 2 led to a temperature of 1550 ± 20 K, a surface gravity $\log(g/g_0)$ of 4.0, a metallicity of [Fe/H] = 0.5 (highest value from our grid), and a C/O ratio of 0.41 ± 0.05, for a planet mass of $11.5\,M_{\rm Jup}$, in very good agreement with the result of Section 3. Contraining the fit to Solar metallicity resulted in a very low C/O ratio of 0.3, with similar temperature and surface gravity.

Including the GPI Y, J, and H band data from Chilcote et al. (2017) and allowing for a multiplicative scaling factor between GPI and GRAVITY resulted in a temperature of $T = 1590 \pm 20$ K, with a C/O of 0.43 ± 0.05, for a metallicity of [Fe/H] = 0.5. For reference, the typical multiplicative factors needed to scale the GPI spectra on the ExoREM grid were $\simeq 0.85$ for the Y band, and $\simeq 0.9$ for J and H bands.

The results of these different fits are summarized in Table 3, and the best fit obtained using GRAVITY+GPI and a mass prior is shown in Figure 4.

4.3. Free retrieval with petitRADTRANS

4.3.1. Retrieval forward model

In addition to fitting a model grid to the β Pic b observation, we carried out a free retrieval. To this end, the spectra were compared to the predictions of a spectral synthesis code, where the atmospheric structure was parametrized. In such an approach more weight is given on atmospheric conditions as constrained by the data, while principles such as radiative-convective equilibrium do not have to be strictly fulfilled. This approach was motivated by the work of Line et al. (2015, 2017); Zalesky et al. (2019) for clear, and Burningham et al. (2017) for cloudy brown dwarfs, in which the power of free retrievals to constrain condensation and cloud processes has been demonstrated.

Our "forward model", used for predicting the spectra, was constructed using petitRADTRANS (Mollière et al. 2019). Because the atmosphere of β Pic b is expected to be cloudy, we added scattering to petitRADTRANS. We verified the calculations by comparing to spectra of self-consistent models for cloudy, self-luminous planets obtained with petitCODE (Mollière et al. 2015, 2017), which agreed excellently.

One benefit of using a free retrieval is that one of the most uncertain physical processes, namely the formation of clouds, can be parametrized, letting the observations constrain the cloud mass fraction and particle size distribution. A related approach was taken by Burningham et al. (2017), who carried out free retrievals for cloudy brown dwarfs for the first time. Here, we assume that our clouds consist of iron and silicate particles, which fixes the location of the cloud base for a given temperature profile. The cloud parameterization of Burningham et al. (2017) was even more general. One of them retrieved the cloud location (where it becomes optically thick), scale height, the single scattering albedo, as well as the power law slope of the opacity.

For the fits presented here, we parametrized the clouds using the Ackerman & Marley (2001) cloud model. However, in contrast to the usual treatment in grid models (see, e.g. Ackerman & Marley 2001; Marley et al. 2012; Morley et al. 2014; Mollière et al. 2017; Samland et al. 2017; Charnay et al. 2018), we retrieved all of its three parameters. First, the settling parameter $f_{\rm sed}$, which is the mass-averaged ratio between the settling and mixing velocity of the cloud particles. This determines the decrease of the cloud mass fraction with altitude, which we set to be $\propto P^{f_{\rm sed}}$. Second, the atmospheric mixing coefficient K_{zz}, which sets the average particle size, once $f_{\rm sed}$ is fixed. In grid models using the Ackerman & Marley (2001) cloud model, this parameter is usually fixed by mixing length theory (with overshooting) or held constant. Third, the width
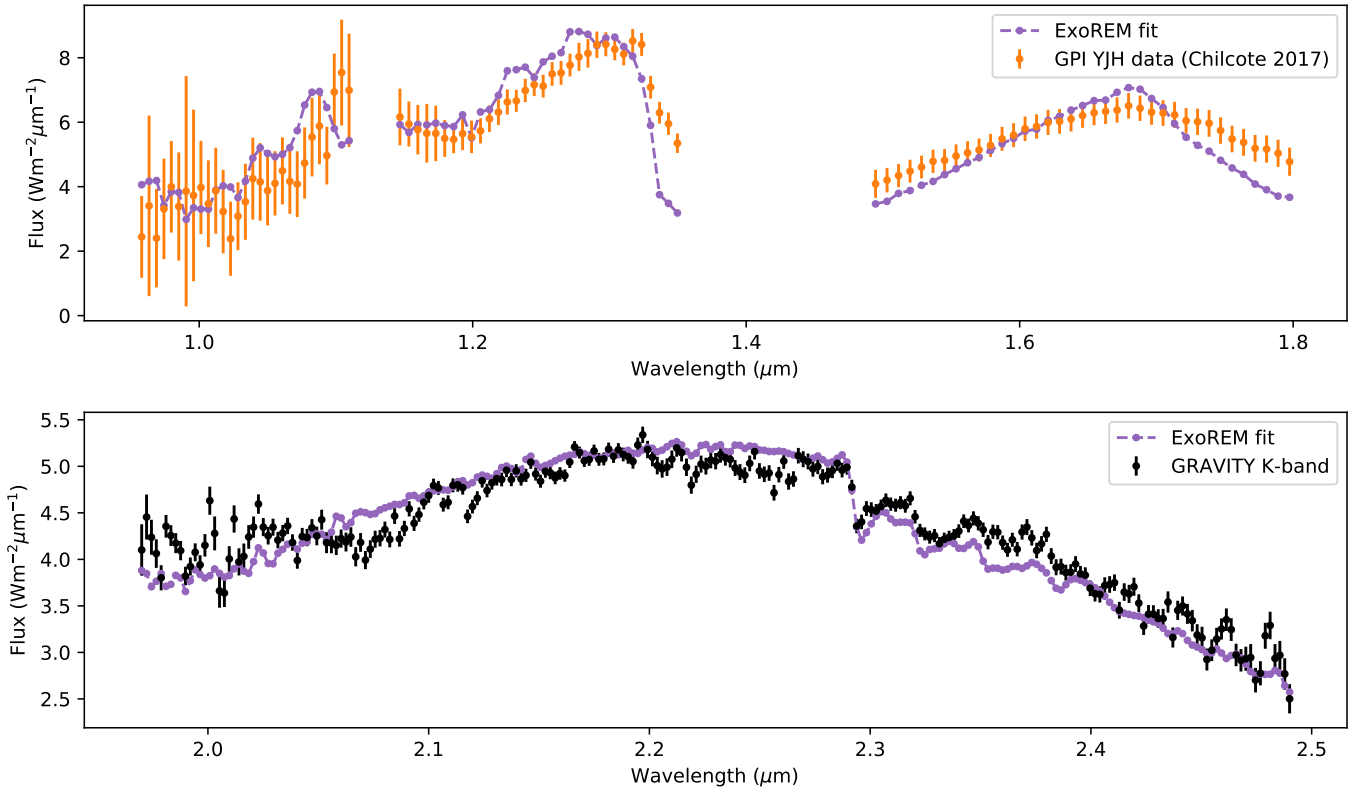
Fig. 4. Best fit obtained with the ExoREM atmospheric model (Charnay et al. 2018) using GPI Y, J, H + GRAVITY K bands, and a mass prior.

of the log-normal particle size distribution σ_g, which is normally also kept constant. The cloud mass fraction at the bottom of the cloud was a free parameter, whereas the position of the cloud base was found by intersecting the P-T profile with the saturation vapor pressure curves (taken from Ackerman & Marley 2001, in the corrected pressure units) of the cloud species we considered, Fe and $MgSiO_3$.

In the future we plan to also test the Burningham et al. (2017) models of clouds, as they are more general, and do not assume the prevalence of a certain cloud species. Moreover, retrieving the power law slope and albedo of the cloud opacities may represent a better choice: for us this is encoded in our choice of cloud species, particle sizes and width of the log-normal particle size distribution, in a non-trivial way. Based on their findings, Burningham et al. (2017) suggest that a log-normal particle size distribution may not be the ideal choice, and that a Hansen distribution (Hansen 1971) may be better.

While carrying out verification retrievals of cloudy petitCODE spectra, we found that we had to be very careful with how the temperature was parametrized. If the temperature model was too flexible (e.g., independent layers + p-spline interpolation, as used in Line et al. 2015), test retrievals of cloudy synthetic spectra lead to clear, hot atmospheres with shallow temperature gradients, that well matched the synthetic input spectrum, but were inconsistent with the input temperature and cloud structure. This could indicate that the cloud-free solutions occupied a larger prior volume, and were thus favored when using a Markov Chain Monte Carlo (MCMC) retrieval.

Specifically, we found it to be necessary to impose a temperature profile in the photospheric region that follows the Eddington approximation, that is

$$T_{\mathrm{phot}}^4 = \frac{3}{4}T_0^4\left(\frac{2}{3}+\tau\right), \qquad (3)$$

where T_0 is normally the internal temperature (taken to be a free nuisance parameter here) and τ the optical depth. This shape was used from $\tau = 0.1$ to the radiative-convective boundary, below which we forced the atmosphere onto a moist adiabat. The optical depth was modeled via

$$\tau = \delta P^\alpha, \qquad (4)$$

where δ and α are free parameters. A quite strict prior was imposed on α. We rejected all models where $|\alpha - \tilde{\alpha}| > 0.1$, where $\tilde{\alpha}$ is the power law index measured from the opacity structure of a given forward model realization. It was obtained from estimating the Rosseland mean opacity using the non-gray opacity of the atmosphere, across the spectral range of the observations. These altitude-dependent values were then used to calculate an optical depth $\tilde{\tau}$, and from this

$$\tilde{\alpha} = \left\langle \frac{d\log\tilde{\tau}}{d\log P} \right\rangle. \qquad (5)$$

Here $\langle \rangle$ denotes the average over the photospheric region. This prior ensures that the parametrized, pressure-dependent opacity is consistent with the atmosphere's non-gray opacity structure. In future applications of the

Fit performed	T (K)	$\log(g/g_0)$	metallicity [Fe/H]	C/O ratio	Mass ($M_{\rm Jup}$)	$\chi^2_{\rm red}$
ExoREM						
GRAVITY data only	1700 ± 50	3.5	-0.5	≤ 0.30	2.0	3.4
GRAVITY + GPI YJH band data	1590 ± 20	4.0	0.5	0.43 ± 0.05	$12.4^{(*)}$	2.4
petitRADTRANS						
GRAVITY data only	1847 ± 55	$3.3^{+0.54}_{-0.42}$	$-0.53^{+0.28}_{-0.34}$	$0.35^{+0.07}_{-0.09}$	$1.4^{+3.94}_{-0.87}$	$2.6^{(a)}$
GRAVITY + GPI YJH band data	1742 ± 10	$4.34^{+0.08}_{-0.09}$	$0.68^{+0.11}_{-0.08}$	$0.43^{+0.04}_{-0.03}$	$15.43^{+2.91}_{-2.79}$	$2.1^{(b)}$

Table 3. Results obtained with the ExoREM model grid and free parameter retrieval petitRADTRANS. (*) Using a mass prior in the fit. (a) Mean value of 100 posterior samples, assuming 17 free parameters, using the GRAVITY covariance matrix. (b) Mean value of 100 posterior samples, assuming 21 free parameters, using the GRAVITY covariance matrix.

parametrized P-T we will test to not downright reject models with too large $|\alpha - \tilde{\alpha}|$. Instead one could adapt the log-likelihood by adding

$$L_\alpha = -\frac{(\alpha - \tilde{\alpha})^2}{2\sigma_\alpha^2} - \frac{1}{2}\log\left(2\pi\sigma_\alpha^2\right) \qquad (6)$$

and fitting for σ_α as a free parameter. Moreover, other P-T parametrizations, for example that of Madhusudhan & Seager (2009), should be tested. This parametrization was also used in Burningham et al. (2017).

In order to prevent the location of the Eddington photosphere to be unrealistically deep in the atmosphere, we also rejected models where

$$P(\tau = 1) > 5P(\tilde{\tau} = 1). \qquad (7)$$

Above the photosphere the temperature was freely variable. We modeled these high altitudes by retrieving the temperature of three locations spaced equidistantly in $\log(P)$ space, and spline interpolating between them.

The chemical abundances and moist adiabat of the atmosphere were found by interpolating in a chemical equilibrium table which contained these quantities as a function of T, P, C/O and [Fe/H]. This table was calculated with the equilibrium chemistry code described in (Mollière et al. 2017). In addition, we also retrieved a quench pressure $P_{\rm quench}$. At pressures smaller than $P_{\rm quench}$ the abundances of CH_4, H_2O and CO were held constant, so as to model the effect of chemical quenching in regions where the chemical reaction timescales become longer than the mixing timescales (see, e.g., Zahnle & Marley 2014).

The following absorption opacity sources where included: CO, H_2O, CH_4, NH_3, CO_2, H_2S, Na, K, PH_3, FeH, VO, TiO, H_2-H_2 (CIA), H_2-He (CIA), Fe clouds (crystalline particles, irregularly shaped), $MgSiO_3$ clouds (crystalline particles, irregularly shaped). The following scattering opacity sources where included: H_2 Rayleigh scattering, He Rayleigh scattering, Fe clouds, $MgSiO_3$ clouds. The opacity references can be found in Mollière et al. (2019).

Using the setup described above, we were able to successfully retrieve the spectrum and atmospheric parameters for a synthetic observation of a cloudy, self-consistent model obtained with petitCODE. The implementation of the retrieval forward model presented here will be described in detail in an upcoming paper. It will contain a description of how the scattering was added, and the verification thereof, as well as the verification retrieval test.

The parameter estimation was carried out using emcee (Foreman-Mackey et al. 2013). Due to the complex priors resulting from the temperature parametrization, the high dimensionality, and a potentially multimodal posterior, the acceptance fraction is low (of the order of 1-2 %) such that one million models were drawn (started around the best-fit position of a pre-burn) to obtain results where the walker positions had converged. As is common for all parameter estimations using an MCMC method, we cannot guarantee that the retrieval results have converged to the true global maximum of the log-probability. While the multi-modality of our model is an inherent property, the acceptance fraction can be improved by setting up the chain closely around the best-fit position of the pre-burn-in run (Foreman-Mackey et al. 2013). In addition, we are currently working on implementing a parameter estimation using nested sampling, which also applies a clustering algorithm for the parameter estimation. This should alleviate the low acceptance rate problem, and lead to a complete sampling of the parameter space (Feroz & Hobson 2008; Feroz et al. 2009). With these current limitations in mind, we note that we retrieved similar values as in the grid retrieval with Exo-REM in Section 4.2, and could successfully retrieve self-consistent cloudy input models when testing our method.

4.3.2. Retrieval parameter results

Our forward model has 17 free parameters: 6 for the temperature model described above, 3 for the abundances (C/O, [Fe/H], $P_{\rm quench}$), 5 for the clouds (the cloud mass fraction of the $MgSiO_3$ and Fe at the cloud base, the settling parameter $f_{\rm sed}$, the eddy diffusion coefficient K_{zz}, the width of the log-normal particle size distribution σ_g), the gravity $\log(g)$, the planetary radius $R_{\rm Pl}$, and the abundance of FeH (currently not included in the chemical table). We used uniform or log-uniform priors for all parameters. In addition to the parameters above we allowed for an individual scaling of the GPI (Y, J, H) bands by up to ± 50 %, and by up to ± 2.5 % for the GRAVITY data. A large value for the scaling of the GPI data was chosen because a similar scaling value was found when comparing the GPI and SPHERE J-band for 51 Eri b in Samland et al. (2017). The maximum scaling we retrieve for β Pic b is 13 % in the GPI Y-band, see below.

Similar to the Exo-REM analysis in Section 4.2, we ran retrievals for a GRAVITY only and GRAVITY + GPI case. In Figure 5 we show the results when fitting the GRAVITY and GPI data together, but without imposing a prior on the mass of β Pic b. For producing this plot, we sampled the posterior distribution 100 times, and plot both the spectra and the accordingly scaled data points. We give the me-
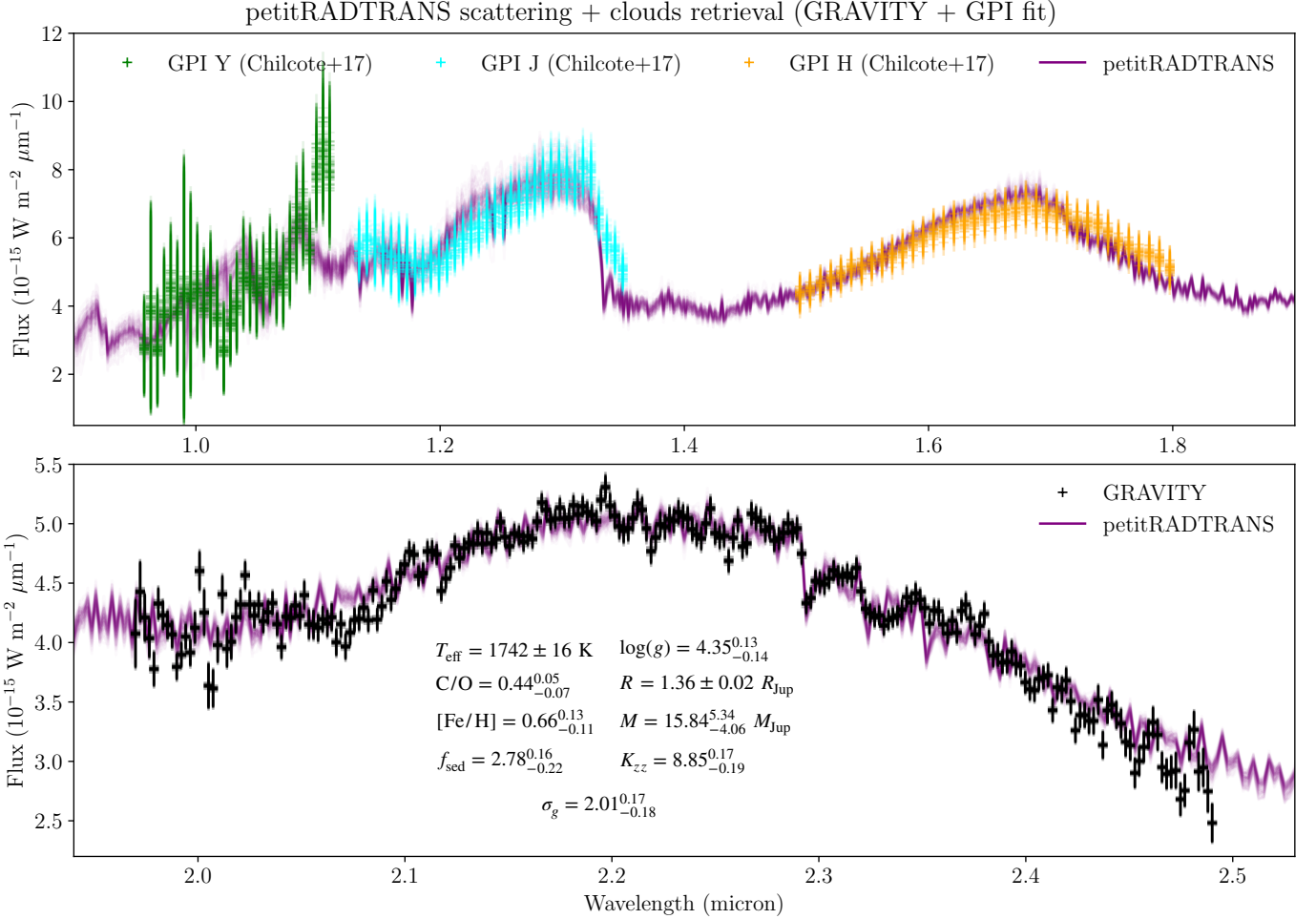
Fig. 5. Results of the combined (GRAVITY+GPI) fit of the β Pic b spectrum with petitRADTRANS. No prior on the mass was used in the fit, and the spectroscopically retrieved mass is consistent with the astrometric value. For producing this plot, 100 samples were drawn from the posterior distribution, for both the model and the data scaling. The 2-d projection of the posterior can be found in Appendix B. *Top panel:* the GPI Y, J and H-band data of Chilcote et al. (2017) are plotted as green, cyan, and orange points with error bars, respectively, the petitRADTRANS models are plotted as purple solid lines. The fit is dominated by the high S/N of the GRAVITY data (shown in the *bottom panel*), leading to a worse fit in the GPI bands, see text. *Right panel:* the GRAVITY data are shown as black points with errorbars, the petitRADTRANS models are plotted as purple solid lines.

dian, and 16 and 84 percentile values of some of the free parameters in the Figure. The full posterior and resulting temperature confidence envelopes are shown in Appendix B. The effective temperature was obtained from integrating the flux of the sampled spectra from 0.5 to 20 microns. The mass was calculated from the $\log(g)$ and R_P values of the posterior sample.

As can be seen in Figure 5, the GRAVITY data can be well fit. At least two CO bandheads at ~ 2.3 micron are visible in the data. The GPI data is less well fit, which may be partially due to the high S/N of the GRAVITY data, which dominates the fit. We found that the fit of the GPI data improved when increasing the error bars of the GRAVITY data. Likewise, we found that the slope in the red part of the GRAVITY spectrum can be fit better if the GPI data are neglected. The retrieved parameters are presented in Table 3, together with the ExoRem results for comparison. Most interestingly, petitRADTRANS retrieves a mass which is consistent with the values from the astrometric measurement in Section 3, without the need of imposing

a prior on the mass. Here we find $M_P = 15.43^{+2.91}_{-2.79}\,M_{\mathrm{Jup}}$, which is consistent with the values presented in Table 2.

From the retrievals presented here, it appears that the GRAVITY K-band data are useful for constraining the planetary C/O ratio, while adding the GPI Y, J, and H-bands is important for obtaining better constraints on the planet's gravity, and hence mass (petitRADTRANS retrieves a too low mass if the GPI data are neglected, see Section 4.4 below). This is consistent with the sensitivity of the NIR YJH bands to gravity (atomic and molecular features, as well as the H band shape), while the CO absorption in the K band is not expected to probe the surface gravity very well (see, e.g., Béjar & Martín 2018). As demonstrated here, estimating the planetary mass from a planet's spectrum alone may become feasible when applying free retrievals over a broad spectral range, such as carried out here with petitRADTRANS.

4.4. Comparison between the grid and free retrieval

In agreement with the Exo-REM fit for GRAVITY+GPI, petitRADTRANS obtains a cloudy atmosphere, and a slightly sub-solar C/O of $0.43^{+0.04}_{-0.03}$ (Exo-REM found 0.43 ± 0.05). The free retrieval obtains a metallicity of $0.68^{+0.11}_{-0.08}$. This is higher than Exo-REM, where 0.5 was found, but this was at the boundaries of the Exo-REM grid, and could likely be higher. This could also be the reason for the slightly higher $\log(g/g_0)$ value ($4.34^{+0.08}_{-0.09}$, and 4 for Exo-REM), due to the gravity-metallicity correlation. petitRADTRANS finds an effective temperature which is higher than in the Exo-REM fit by about 150 K (1742 ± 10 K, compared to 1590 K for Exo-REM). The larger radius found by Exo-REM is likely due to the lower temperature it retrieved, so as to conserve the total amount of flux. At the estimated age of β Pic (24 ± 3 Myr; see Bell et al. 2015), a radius of 1.7 $R_{\rm Jup}$ (the value Exo-REM retrieved) requires masses in excess of 20 $M_{\rm Jup}$, and effective temperatures of around 2500 K, when considering hot start models (Spiegel et al. 2011). Core accretion models under the warm[2] start assumption, which include deuterium burning, require similarly large masses and temperatures, but potentially somewhat younger ages (Mollière & Mordasini 2012; Mordasini et al. 2017), and would put the planet firmly into the mass regime of brown dwarfs currently undergoing deuterium burning. Hence, the large radius retrieved by Exo-REM is likely to be inconsistent with the retrieved mass and temperature. The values of the mass, temperature, and radius (1.36 $R_{\rm Jup}$) retrieved by petitRADTRANS agree with both the cold and hot start predictions, given the age of β Pic.

Also when fitting only the GRAVITY data, the petitRADTRANS results are mostly consistent with Exo-REM. Without a mass prior we find C/O $= 0.35^{+0.07}_{-0.09}$ (Exo-REM found $\lesssim 0.3$), [Fe/H] $= -0.53^{+0.28}_{-0.34}$ (Exo-REM found -0.5), $\log(g/g_0) = 3.3^{+0.54}_{-0.42}$ (Exo-REM found 3.5), $M = 1.4^{+3.94}_{-0.87}$ $M_{\rm Jup}$ (Exo-REM found 2 $M_{\rm Jup}$). Only the temperature is larger again, at 1847 ± 55 K (petitRADTRANS), compared to 1700 K (Exo-REM).

In summary, a free retrieval approach gives similar results to a more classical retrieval from a grid of forward models. We note that here a free retrieval appears to lead to physically more consistent results when constraining radii and effective temperatures. Another possible cause for the differences could be how the opacities of gas and clouds are treated. For the gas opacities we note that petitRADTRANS uses the opacity database of petitCODE, the latter of which was successfully benchmarked with Exo-REM in Baudino et al. (2017). Small remaining differences, identified to stem from the use of different line lists in Baudino et al. (2017), have since been removed by updating the opacity database of petitCODE/petitRADTRANS in 2017.

4.5. Comparison to Chilcote et al. (2017)

A substantial analysis of the NIR spectrocopy of β Pic b was carried out in Chilcote et al. (2017), using GPI YJHK band spectra. The data were compared to low gravity and field brown dwarf spectra, the derived bolometric luminosity was compared to evolutionary models, and spectral fits were carried out with four different model grids.

Comparing their bolometric luminosity to evolutionary models (hot start models of Baraffe et al. 2003), Chilcote et al. (2017) found a mass of 12.9 ± 0.2 $M_{\rm Jup}$, an effective temperature of 1724 ± 15 K, a surface gravity of $\log(g/g_0) = 4.18 \pm 0.01$ and a radius of 1.46 ± 0.01 $R_{\rm Jup}$. Their mass measurement is consistent both with our astrometrically and spectroscopically inferred mass values. Moreover, the other values inferred from the the YJHK fit of petitRADTRANS are close to the evolutionary values of Chilcote et al. (2017), but not within the uncertainties of one another (e.g., 1742 ± 10 K vs. 1724 ± 15 K). As noted in Chilcote et al. (2017), these uncertainties do not contain a contribution of the model uncertainties, and the true uncertainties must be larger. The same holds for our retrievals and fits carried out here. The Exo-REM fits with mass prior lead to a similar agreement in gravity, but the radii and temperatures are further away from the Chilcote et al. (2017) values, with the planet being cooler, and thus larger, in the Exo-REM fits.

The best grid model fit of the combined photometry and spectroscopy in Chilcote et al. (2017) was obtained with Drift-PHOENIX (e.g., Helling et al. 2008), where $T_{\rm eff} = 1651$ K, $\log(g) = 3$ and $R = 1.58$ $R_{\rm Jup}$ was found, leading to a mass of ~ 1 $M_{\rm Jup}$. The AMES-Dusty (e.g., Allard et al. 2001) fit gave the highest mass (most consistent with our astrometric, spectroscopic and Chilcote et al. 2017's evolutionary mass), namely 17 $M_{\rm Jup}$. The AMES-Dusty best-fit values are also closer to the evolutionary parameters derived in Chilcote et al. (2017), at $T_{\rm eff} = 1706$ K, $\log(g) = 4.5$ and $R = 1.18$ $R_{\rm Jup}$, but at an overall worse χ^2 than the Drift-PHOENIX fit.

In summary, the results of our spectral characterization compare well with the evolutionary values inferred for β Pic b in Chilcote et al. (2017). Especially the parameter values of the free retrieval carried out with petitRADTRANS are close to the evolutionary values. It is also noteworthy that our masses, inferred indepently with astrometry or spectral retrieval, are consistent with the evolutionary mass of Chilcote et al. (2017).

5. C/O ratio and the formation of beta Pic b

5.1. Stellar and planetary C/O ratio

Holweger et al. (1997) have shown that the abundances of several elements (C, Ca, Ti, Cr, Fe, Sr, Ba) on the surface of β Pictoris are Solar. But measuring the abundance of oxygen in stars is a notoriously difficult task, due to line blending, deviations from local thermal equilibrium predictions, or sensitivity to the 3D temperature structure of the star (Asplund 2005). As a consequence, to our knowledge, the abundance of oxygen – and hence the C/O ratio – has not yet been reported in the literature. We note, though, that a subsolar C/O ratio (i.e., $\simeq 0.4$) would invalidate most of the following discussion.

In our atmosphere analysis, both the ExoREM grid model fitting and the petitRADTRANS free retrieval point to the same result: the C/O number ratio in β Pictoris b is $\simeq 0.43 \pm 0.05$, which is subsolar (the solar C/O ratio is 0.55, see Asplund et al., 2009).

A number of studies have been done in the recent years to try to find links between planetary formation processes

---

[2] These models are somewhat warmer than the classical cold start assumption (Marley et al. 2007), because the planetesimal accretion is not shut off after the isolation mass is reached.

Species	$n_{\mathrm{species}}/n_{\mathrm{H_2O}}$
H_2O	1
CO	1.67
CO_2	0.33
C (grains)	0.67
O (silicates)	1.54

Table 4. Relative abundances of the different species taken from Table 1 of Öberg et al. (2011), and used in our young β Pic protoplanetary disk. All values are given relative to H_2O.

and element abundances. In particular, Öberg et al. (2011) first attempted to relate the C/O ratio to the position of the different icelines in a protoplanetary system, and to the proportion of gas and solid material accreted by a young planet. They concluded that substellar C/O ratio was a sign of a formation by either gravitational collapse or core-accretion, followed by icy planetesimal enrichment. The objective of this section is to show that the C/O ratio can possibly be used to disentangle the two formation scenarios.

5.2. General model for the evolution of the C/O ratio

In a similar fashion as to Öberg et al. (2011), we assume that the main sources of carbon and oxygen in the protoplanetary disk in which β Pic b formed were CO, CO_2, H_2O, silicates and carbon grains. Assuming a solar C/O ratio for the star, the table of relative abundances given in Öberg et al. (2011) is valid, and we use it as a baseline to set the abundances of each species (see Table 4).

In the framework developed by Öberg et al. (2011), the C/O ratio in the atmosphere of a planet can be calculated from the amount of solid and gaseous material entering its composition. We denote $n_{X,s}$ (resp. $n_{X,g}$) the abundance of element X in the solid phase (resp. gas phase) of the disk, given in number of atoms per unit of disk mass. We also write M_{solid} (resp. M_{gas}) the total mass of solid (resp. gas) entering the composition of the atmosphere of the planet, and $f_{\mathrm{s/g}}$ the dust-to-gas fraction in the disk, which we assume to be equal to 0.01. With these notations, the total number of elements X in the atmosphere of the planet is given by:

$$N_X = \frac{n_{X,s}}{f_{\mathrm{s/g}}} \times M_{\mathrm{solid}} + \frac{n_{X,g}}{1 - f_{\mathrm{s/g}}} \times M_{\mathrm{gas}} \quad (8)$$

And the C/O number ratio is then:

$$\mathrm{C/O} = \frac{n_{C,s}f_{\mathrm{s/g}}^{-1}M_{\mathrm{solid}} + n_{C,g}(1 - f_{\mathrm{s/g}})^{-1}M_{\mathrm{gas}}}{n_{O,s}f_{\mathrm{s/g}}^{-1}M_{\mathrm{solid}} + n_{O,g}(1 - f_{\mathrm{s/g}})^{-1}M_{\mathrm{gas}}} \quad (9)$$

Note that both the numerator and the denominator can be given relative to a reference species without affecting the validity of this Eq. (9). In Table 4 and in all the following, we implictly use abundances relative to H_2O.

The exact values of $n_{C,s}$, $n_{C,g}$, $n_{O,s}$, and $n_{O,g}$ depends on the abundances given in Table 4, and on the state (solid or gaseous) of each species. and hence on the location of the forming planet with respect to the different icelines.

Using ALMA observations, Qi et al. (2015) have shown that the CO iceline in the disk around HD 163296 was likely

to be located at $\simeq$ 90 AU from the star. Other observations of the same system, also performed with ALMA, led Notsu et al. (2019) to conclude that the water iceline was located at a distance of $\leq$ 20 AU. Since HD 163296 is also an A-type star, these two values give an idea of the possible location of the H_2O and CO icelines in the β Pic system. However, little is known about the relationship between the current orbit of β Pic b and its exact formation location, and about possible variations of the locations of these icelines between systems. Thus, no definitive assumption can be made as to where the planet formed in comparison to the water iceline, and the two options must be considered: a formation within the water iceline, and a formation between the water and the CO_2 icelines.

From there, the terms $n_{C,s}, n_{C,g}, n_{O,s}$, and $n_{O,g}$ from Eq. (9) can be determined from the values listed in Table 4. For a planet forming within the water iceline, we have:

$$\begin{cases} n_{O,g} = n_{\mathrm{H_2O}} + n_{\mathrm{CO}} + 2 \times n_{\mathrm{CO_2}} = 3.33 \\ n_{O,s} = n_{O \text{ (silicates)}} = 2.12 \\ n_{C,g} = n_{\mathrm{CO}} + n_{\mathrm{CO_2}} = 2.0 \\ n_{C,s} = n_{C \text{ (grains)}} = 0.67 \end{cases} \quad (10)$$

And for a planet forming between the water and CO_2 icelines:

$$\begin{cases} n_{O,g} = n_{\mathrm{CO}} + 2 \times n_{\mathrm{CO_2}} = 2.33 \\ n_{O,s} = n_{\mathrm{H_2O}} + n_{O \text{ (silicates)}} = 3.12 \\ n_{C,g} = n_{\mathrm{CO}} + n_{\mathrm{CO_2}} = 2.0 \\ n_{C,s} = n_{C \text{ (grains)}} = 0.67 \end{cases} \quad (11)$$

5.3. C/O ratio in the gravitational collapse paradigm

We consider the case of a formation through gravitational collapse (Bodenheimer 1974), a violent mechanism which shares similarities with star formation. In this scenario, an entire region of the circumstellar disk becomes unstable, and rapidly collapses to form a protoplanet, which then slowly contracts and cools down.

The total mass of solid entering in the composition of the atmosphere of a planet formed through gravitational collapse can be separated in two terms: the mass of solid initially contained in the disk fragment which collapsed to create the protoplanet, and the mass of solid planetesimals later accreted by the protoplanet. The solid mass contained in the initial clump is directly related to the dust-to-gas ratio of the disk, and we can write:

$$M_{\mathrm{solid}} = f_{\mathrm{s/g}}M_{\mathrm{planet}} + M_{\mathrm{accreted}} \quad (12)$$

This equation assumes that no core has formed in the young protoplanet, which, for a planet as massive as β Pic b is reasonable (Helled & Schubert 2008). For a planet less massive, for which a core could form, sedimentation of a fraction of the initial solid mass on the core should be taken into account.

Injecting the definition of M_{solid} into Eq. (9), and using $f_{\mathrm{s/g}} = 0.01$, $M_{\mathrm{planet}} = 12.7\,M_{\mathrm{Jup}}$, as well as the values for C and O abundances given in Eq. (10) or (11), it is possible to determine the C/O ratio as a function of the mass of accreted planetesimals M_{accreted} in the gravitational collapse paradigm. The results is given in Figure 6, for two possible formation locations: within the water iceline, and between the water and carbon dioxide icelines. We have
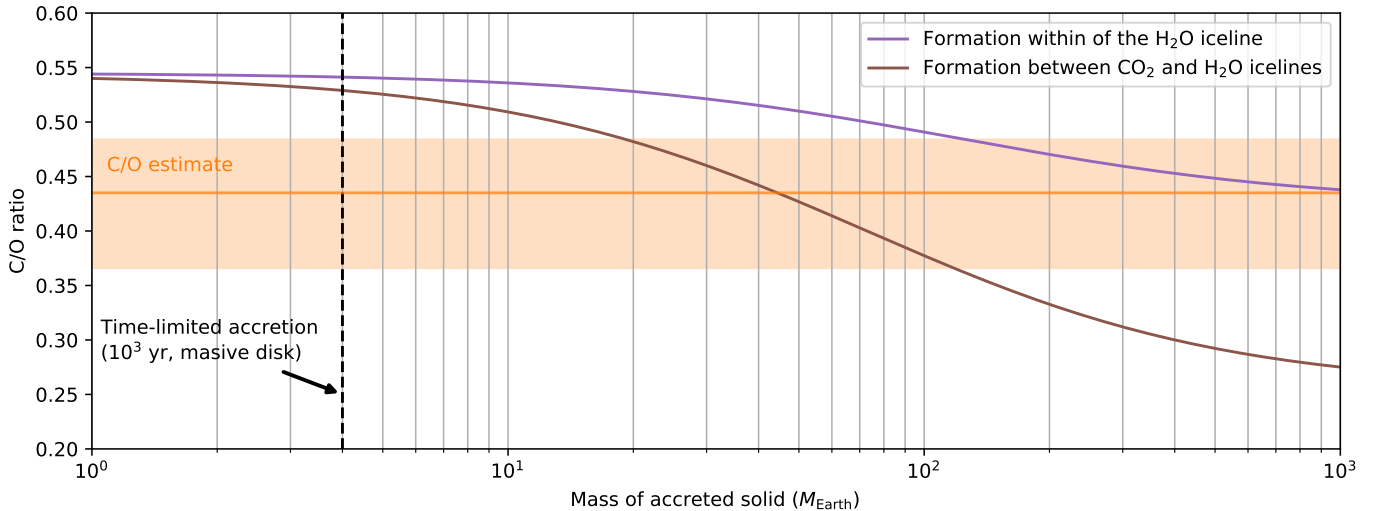
Fig. 6. Gravitational collapse scenario: evolution of the C/O ratio as a function of the total mass of solid accreted after the initial formation of the protoplanet. The purple curve corresponds to a formation within the H_2O iceline, and the brown curve to a formation between the H_2O and CO_2 icelines. The orange area gives the 68% confidence interval for the value of the C/O ratio. Dashed vertical lines corresponds to different solid accretion limits discussed in the text.

also added the $1\,\sigma$ confidence intervals of our ExoREM and petitRADTRANS measurements on this graph. This figure shows that a formation bewteen the H_2O and CO_2 icelines is more favorable to a large deviation from the stellar C/O ratio, mainly due to the injection of oxygen coming from solid water ice during planetesimal accretion.

The formation of a planet by gravitational instability can be separated in a few different steps (Bodenheimer 1974): formation of the initial clump in the disk, quasi-equilibrium contraction, hydrodynamic collapse, and a new hydrostatic quasi-equilibrium phase. Accretion of planetesimals is thought to be efficient only during the pre-collapse phase (Helled & Schubert 2009). The duration of this phase decreases with increasing planet mass, and typical values ranges from a few 10^5 years for a Jupiter mass planet, to less than 10^3 years for more massive planets (Decampli & Cameron 1979; Bodenheimer et al. 1980). Using the model proposed by Helled & Schubert (2009), the mass of planetesimal accreted during the pre-collapse phase of β Pic b can be estimated using:

$$M_{\mathrm{accreted}} = \int_{0}^{t_{\mathrm{collapse}}} \pi R_{\mathrm{capture}}^2(t)\sigma(a,t)\Omega(a)\mathrm{d}t \qquad (13)$$

Where t_{collapse} is the time of collapse, R_{capture} the protoplanet's capture radius, σ the surface density of solids in the disk at the location of the protoplanet, and Ω the orbital frequency.

Andrews & Williams (2005) presented a large survey of 153 young stellar objects in the Taurus-Auriga star forming region. Among all these objects, AB Aur and V892 Tau are two A-type stars, for which they give an estimate of the mass: 0.004 $M_{\odot}$ and 0.009 $M_{\odot}$. Considering all stellar types, the median disk-to-star mass ratio they found is 0.5%. More recent studies of protoplanetray disk demographics based on ALMA observations yielded similar results, with typical dust to star mass ratios of $\simeq 10^{-4.5}$ (Pascucci et al. 2016; Ansdell et al. 2017), i.e. disk-to-star mass ratios of $\simeq 0.3\%$, assuming a dust-to-gas ratio of 1%.

Considering the upper limit of an extremely massive disk ($M_{\mathrm{disk}} = 0.1\ M_{\odot}$), and using a power-law for the surface density ($\sigma = \sigma_0\,(a/5\ \mathrm{AU})^{-\alpha}$, with $\alpha = 1.00$), the solid density at $a = 11$ AU is:

$$\sigma(11\ \mathrm{AU}) \simeq 6\ \mathrm{g/cm^2} \qquad (14)$$

The orbital period of the planet is ~ 20 yr (Wang et al. 2016; Lagrange et al. 2018, Section 3 of this work). The capture radius decreases with the contraction of the planet, but an optimistic value would be 2 to 3×10^{12} cm for a 1 M_{Jup} planet (Helled et al. 2006). For a planet 10 times more massive, the effective radius could be $\simeq 5 \times 10^{12}$ cm. This yields:

$$M_{\mathrm{accreted}} \simeq 4 \times M_{\mathrm{Earth}} \times \frac{t_{\mathrm{collapse}}}{1000\ \mathrm{yr}} \qquad (15)$$

The corresponding accretion limit has been added to Figure 6, for a reasonable assumption of $t_{\mathrm{collapse}} = 10^3$ yr.

Taking into account the effective time available for efficient planetesimal accretion during the pre-collapse stage, the low C/O ratio measured with GRAVITY is difficult to explaine, even in the case of a planet forming oustide the H_2O iceline. For the C/O ratio to reach a value of $\simeq 0.43$, we need to assume a massive protoplanetary disk and an unusually long time for the pre-collapse phase, or an extremely efficient accretion (with an accretion rate of $4 \times 10^{-3}\ M_{\mathrm{Earth}}/\mathrm{yr}$).

5.4. C/O ratio in the core-accretion paradigm

Core-accretion is another formation mechanism, in which an initial solid core forms, and slowly accretes gas from the disk. When the mass of gas is roughly the same as the mass of the core, the protoplanet enters a phase of "runaway gas accretion", during which it gains significant amount of gas over a short time (Lissauer & Stevenson 2007). In this scenario, the formation of a planet is a much longer process
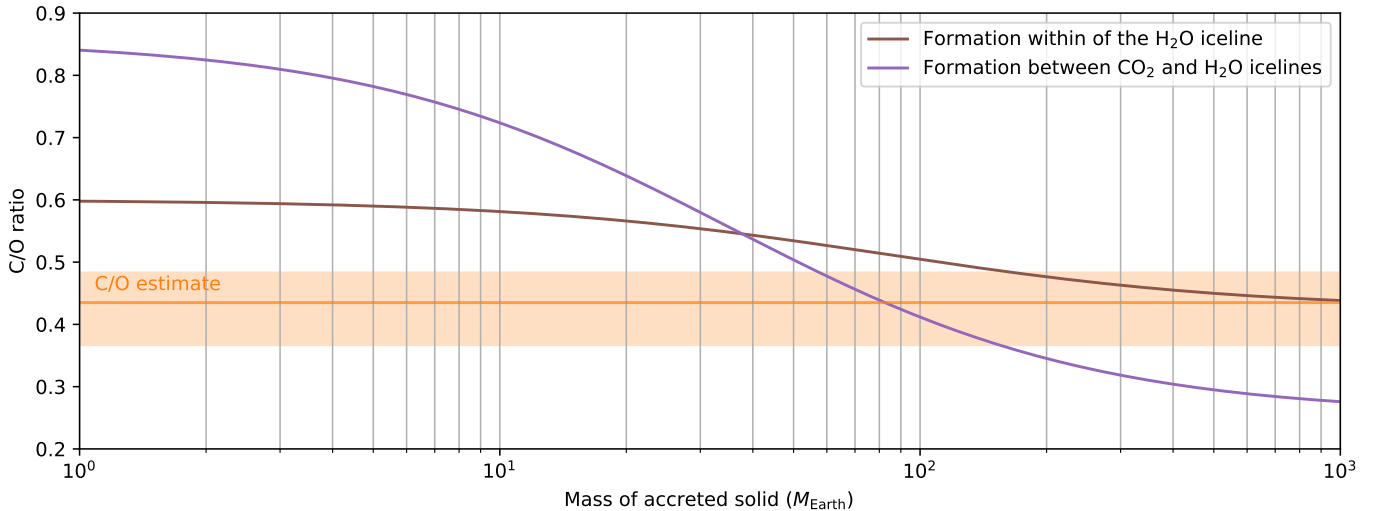
Fig. 7. Core-accretion scenario: evolution of the C/O ratio in the atmosphere of β Pic b as a function of the total mass of solid accreted by the protoplanet, for a formation between the H_2O and the CO_2 icelines, or within the H_2O iceline. The orange area gives the 68% confidence interval for the value of the C/O ratio.

than with gravitational instability, which gives more time to enrich the proto-atmosphere in solid material and to lower its C/O ratio.

Mordasini et al. (2016) explored the effect of planetesimal enrichment coupled with disk composition, in a core-accretion scenario. They focused on the case of Jupiter mass planets migrating to short period orbits ("hot Jupiters"), which is a different archetype than β Pic b. But the general sequence of events they use to form their planets in the core-accretion paradigm can still be applied to β Pic b, only leaving out the inward migration part. First, the core of the planet forms from the accretion of solid material. Then, once the core has formed, the protoplanet starts accreting a gaseous envelope which, during its formation, is enriched by the accretion of disintegrating planetesimals. When the planet reaches a critical mass, runaway accretion occurs, and the mass of the planet significantly increases. This runaway gas accretion clears a gap in the disk, and ends the formation of the planet.

In the gravitational instability scenario, because the formation of the planet happens so quickly compared to typical timescales of disk evolution, the gas and solid making the atmosphere necessarily have a stellar combined C/O. If the solid and gas in the atmosphere are in the same proportion as they are in the disk ($M_{solid} = f_{s/g}M_{gas}$), the C/O of the atmosphere is stellar. A deviation of the solid to gas proportion in the atmosphere is required to alter the C/O ratio.

In the case of core-accretion, the situation is different. Witout planetismal enrichment before the runaway gas accretion phase, the atmosphere of the planet would not be made of a mixture of gas and solid material, but purely of gas. Thus, without planetesimal enrichment, the atmospheric C/O ratio in the core-accretion paradigm can be expected to be close to the C/O ratio of the gas in the disk, that is, superstellar.

In this core-accretion paradigm, it is still possible to use Eq. (9) to calculate how the final C/O ratio of the atmosphere is impacted by the mass of solid material accreted before the runaway accretion phase. But in this case, all of the solid mass M_{solid} corresponds to accreted material: $M_{solid} = M_{accreted}$, as opposed to Eq. (12).

In Figure 7, we show the evolution of the C/O ratio as a function of the mass of accreted planetesimals, for a formation thourgh core-accretion, within the water iceline, or between the water and CO_2 icelines. In this scenario, it is possible to reach C/O values compatible with our GRAVITY measurement with accretion of $\simeq 80\ M_{Earth}$, if the planet formed between the water and CO_2 icelines. A formation within the water iceline is more diffcult to explain, as it would require at least $150\ M_{Earth}$ of solid material enrichment to reach the upper limit of the $1\,\sigma$ interval on the C/O measurement, and up to several $10^2\ M_{Earth}$ to reach a value of 0.43.

6. Summary and conclusions

In this work, we presented the first VLTI/GRAVITY spectro-interferometric observation of the giant planet β Pictoris b. Using an adequate data reduction technique detailed in the appendix of this paper, we extracted a high quality K-band spectrum of the planet, at a resolution of $R = 500$. We also derived the most precise relative astrometry obtained to date on this object, with an error of $\simeq 40\ \mu$as.

We find that the astrometry disfavors circular orbits for β Pic b, with a value of $e \simeq 0.15^{+0.05}_{-0.04}$. It remains unclear how a massive planet like β Pic b can acquire such a significant eccentricity. Using this new astrometric datapoint together with previously published visual astrometry and Hipparcos/Gaia data, we were able to derive an estimate of the dynamical mass of β Pictoris b, in a similar fashion as to what Snellen & Brown (2018) and Dupuy et al. (2019) did. Our value is compatible with these previous studies, with a best estimate of $12.7 \pm 2.2\ M_{Jup}$.

We were also able to retrieve a similar mass, albeit with larger error bars, using only the spectral data. Using a free retrieval, including the effect of scattering and clouds, with petitRADTRANS (Mollière & Snellen 2019) to fit the spectrum of β Pic b in Y, J, H, and K bands (Y, J, H from Chilcote et al. 2017, K from this work), we obtained a mass of $15.43^{+2.91}_{-2.79}\,M_{\mathrm{Jup}}$. This constitutes a rare case of validation of an atmospheric model with a model-independent measurement.

We performed an in-depth analysis of the K-band spectrum extracted from our GRAVITY observation using two different approaches: forward modeling with the ExoREM code (Charnay et al. 2018), and free-retrieval with petitRADTRANS. We found that both approaches point to a C/O ratio of C/O $= 0.43 \pm 0.05$.

We showed that, if the C/O ratio of the host star β Pictoris is Solar, it is difficult to explain this C/O ratio with a gravitational collapse formation scenario. This is mainly due to the high mass of β Pictoris b, which has the dual consequence of requiring large amount of planetesimal enrichment to lower the initial C/O ratio, while at the same time making the whole formation process extremely short. In this case, it appears that a slower formation via core-accretion, somewhere between the H_2O and CO_2 icelines, is more likely. This scenario can potentially explain the sub-solar C/O ratio if the planet was enriched in oxygen by icy planetesimal accretion.

The high metal enrichment we retrieve from the spectral fits appears to corroborate this assessment, with the exact value being quite high and at the edge of what is expected from classical core accretion Mordasini et al. (2016).

This model still comes with several important limitations. One of them is that the exact compositition of the initial protoplanetary disk around β Pic remains largely unknown. Another major issue is the efficiency of the planetesimal enrichment, which we have assumed to be of 100% (i.e., all the solid material accreted by the planet is disintegrated in the atmosphere). This is unlikely to be the case, as fraction of this material can be deposited into the planetary core, or can stay at the bottom of the atmosphere. This is particularly true for the core-accretion scenario, in which the solid material is accreted before most of the gas (Mordasini et al. 2016). Strong vertical mixing can potentially mitigate this problem, but further studies are required to be able to take into account these phenomena. Finally, disk chemistry may also play a role. For example, Eistrup et al. (2018) have shown that a large fraction of water molecules can be transformed into dioxygen (O_2) over a few Myr, along a chemical pathway detailed in Walsh et al. (2015). Oustide of the water iceline, such a chemical evolution can potentially deplete the solid material from its oxygen, while enriching the gas.

The observations of β Pictoris b presented in this paper show the potential of long-baseline optical interferometry with VLTI/GRAVITY for exoplanet science. The instrument gives access to medium resolution spectroscopy in K-band and high-precision astrometry, which are both extremely useful to characterise giant exoplanets and to start peering into their formation history.

Appendix A: Reduction of the GRAVITY dataset

A.1. Nomenclature and pipeline errors

The data reduction used to extract the beta Pictoris b signal from the GRAVITY observations makes heavy use of complex linear algebra, complex error formalism, and maximum likelihood estimation. To avoid confusion and mistakes, complex numbers in this appendix are underlined (e.g., $\underline{V}$), whereas real numbers are not (e.g., X).

Most of the GRAVITY data manipulated are quantities which depends on the wavelength λ. These quantities can be represented as vectors of size n_λ (the number of wavelength channels) by concatenating the individual values. These vectors are denoted using a bold font. For example, in the case of the complex visibility obtained on baseline b at time t, we denote:

$$\mathbf{\underline{V}}_{b,t} = \begin{pmatrix} \underline{V}(b,t,\lambda_1) \\ \underline{V}(b,t,\lambda_2) \\ \vdots \\ \underline{V}(b,t,\lambda_{n_\lambda}) \end{pmatrix} \tag{A.1}$$

For a given DIT, it is also possible to concatenate all baselines to create a vector of size $n_b \times n_\lambda$, where $n_b = 6$ is the number of baselines. In this case, the subscript b is dropped:

$$\mathbf{\underline{V}}_t = \begin{pmatrix} \mathbf{\underline{V}}_{b_1,t} \\ \vdots \\ \mathbf{\underline{V}}_{b_{n_b},t} \end{pmatrix} \tag{A.2}$$

The complex-conjugate of a complex number $\underline{V}$ is denoted $\underline{V}^*$, and the complex-transpose of a vector or matrix $\mathbf{\underline{A}}$ is denoted $\mathbf{\underline{A}}^\dagger$. It is defined by: $\mathbf{\underline{A}}^\dagger = \mathbf{\underline{A}}^{*T}$ where ^{T} is the transpose operator.

All the λ-vectors are understood as elements of a n_λ-dimension complex linear space (i.e. a linear space for which the scalar field is the set of complex numbers $\mathbb{C}$, rather than the set of real numbers $\mathbb{R}$). Adding the natural scalar product operator (i.e. $\langle V_1, V_2 \rangle = V_1^\dagger V_2$) makes this linear space an Euclidean space. This mathematical structure allows for several useful concepts: it is possible to compute othogonal projections, to use projector matrices, to define orthogonal and/or orthonormal basis, etc.

The data set can be subdivided into two parts: the observations taken with the science fiber on the planet, and the observations taken on the star (see observing log in Table 1). On-planet and on-star phase-referenced visibilities are calculated from the coherent fluxes measured by GRAVITY, called VISDATA in the FITS files generated by the pipeline. The VISDATA are complex numbers, affected by noise. The GRAVITY pipeline reports these errors in another set of complex numbers, called VISERR. The real part of VISERR contains the uncertainties on the real part of VISDATA, and the imaginary part of VISERR contains the uncertainty on the imaginary part of VISDATA. These errors do not take into account any possible correlation between different spectral channels, or between the real and imaginary parts of the visibility. To take into account such correlations, it is necessary to use the covariance/pseudo-covariance formalism of complex random variables.

In our data reduction algorithm, the GRAVITY pipeline errors are systematically replaced by an empirical estimate of the covariance and pseudo-covariance matrices of the visibilities. We assume that the noise affecting the measurements does not vary significantly over the individual DITs of a single exposure file (~ 5 min), but can vary from file to file. We also allow for correlations between different spectral channels and/or between different baselines. Under these assumptions, the errors on the coherent fluxes are best represented by a set of n_{EXP} (the number of exposure files) covariance matrices $\mathbf{W}_k$ and n_{EXP} pseudo-covariance matrices $\mathbf{Z}_k$, both of size $n_b \times n_\lambda$, where $n_b = 6$ is the number of baselines and $n_\lambda = 235$ is the number of wavelength channels. The covariance and pseudo-covariance matrices for each exposure file are estimated directly from the DITs sequence:

$$\mathbf{\underline{W}}_k = \frac{1}{n_{\mathrm{DIT}}-1}\left(\sum_{t=1}^{n_{\mathrm{DIT}}} \mathbf{\underline{V}}_t \mathbf{\underline{V}}_t^\dagger - \frac{1}{n_{\mathrm{DIT}}}\left(\sum_{t=1}^{n_{\mathrm{DIT}}} \mathbf{\underline{V}}_t\right)\left(\sum_{t=1}^{n_{\mathrm{DIT}}} \mathbf{\underline{V}}_t\right)^\dagger \right)$$

$$\mathbf{\underline{Z}}_k = \frac{1}{n_{\mathrm{DIT}}-1}\left(\sum_{t=1}^{n_{\mathrm{DIT}}} \mathbf{\underline{V}}_t \mathbf{\underline{V}}_t^T - \frac{1}{n_{\mathrm{DIT}}}\left(\sum_{t=1}^{n_{\mathrm{DIT}}} \mathbf{\underline{V}}_t\right)\left(\sum_{t=1}^{n_{\mathrm{DIT}}} \mathbf{\underline{V}}_t\right)^T \right)$$

where the dummy t runs over the n_{DIT} DITs of the k-th exposure.

The covariance and pseudo-covariance matrices are always related to the covariance of the real and imaginary parts by the following equations:

$$\mathrm{cov}\left(\mathrm{Re}(\mathbf{\underline{V}}), \mathrm{Re}(\mathbf{\underline{V}})\right) = \frac{1}{2}\mathrm{Re}(\mathbf{\underline{W}} + \mathbf{\underline{Z}}) \tag{A.3}$$

$$\mathrm{cov}\left(\mathrm{Im}(\mathbf{\underline{V}}), \mathrm{Im}(\mathbf{\underline{V}})\right) = \frac{1}{2}\mathrm{Re}(\mathbf{\underline{W}} - \mathbf{\underline{Z}}) \tag{A.4}$$

$$\mathrm{cov}\left(\mathrm{Re}(\mathbf{\underline{V}}), \mathrm{Im}(\mathbf{\underline{V}})\right) = \frac{1}{2}\mathrm{Im}(-\mathbf{\underline{W}} + \mathbf{\underline{Z}}) \tag{A.5}$$

$$\mathrm{cov}\left(\mathrm{Im}(\mathbf{\underline{V}}), \mathrm{Re}(\mathbf{\underline{V}})\right) = \frac{1}{2}\mathrm{Im}(\mathbf{\underline{W}} + \mathbf{\underline{Z}}) \tag{A.6}$$

The covariance and pseudo-covariance matrices can be propagated during the data reduction algorithm by using the complex error propagation equations:

$$\mathrm{cov}(\mathbf{\underline{A}}\mathbf{\underline{V}}) = \mathbf{\underline{A}}\mathbf{\underline{W}}\mathbf{\underline{A}}^\dagger \tag{A.7}$$

$$\mathrm{pcov}(\mathbf{\underline{A}}\mathbf{\underline{V}}) = \mathbf{\underline{A}}\mathbf{\underline{Z}}\mathbf{\underline{A}}^T \tag{A.8}$$

with $\mathbf{\underline{A}}$ any complex matrix of appropriate size.

The $\mathbf{\underline{W}}$ and $\mathbf{\underline{Z}}$ matrices can also be used to resolve linear equations involving complex data. In the case of an unknown real parameter vector $\mathbf{X}$, the solution of the linear problem $\mathbf{\underline{V}} = \mathbf{\underline{A}}\mathbf{X}$ (in the sense of maximum likelihood) is:

$$\hat{\mathbf{X}} = \left(\mathrm{Re}\left(\mathbf{\underline{V}}_2^\dagger \mathbf{\underline{W}}_2^{-1} \mathbf{\underline{A}}_2\right) \left[\mathrm{Re}\left(\mathbf{\underline{A}}_2^\dagger \mathbf{\underline{W}}_2^{-1} \mathbf{\underline{A}}_2\right) \right]^{-1} \right)^T \tag{A.9}$$

where

$$\mathbf{\underline{V}}_2 = \begin{pmatrix} \mathbf{\underline{V}} \\ \mathbf{\underline{V}}^* \end{pmatrix} \tag{A.10}$$

$$\mathbf{\underline{W}}_2 = \begin{pmatrix} \mathbf{\underline{W}} & \mathbf{\underline{Z}} \\ \mathbf{\underline{Z}}^\dagger & \mathbf{\underline{W}}^* \end{pmatrix} \tag{A.11}$$

$$\mathbf{\underline{A}}_2 = \begin{pmatrix} \mathbf{\underline{A}} \\ \mathbf{\underline{A}}^* \end{pmatrix} \tag{A.12}$$

For a complete mathematical derivation of Eq. (A.9), we refer the reader to Appendix B of Nowak (2019).

A.2. Pipeline reduction and phase referencing

The initial step uses the pipeline reduction and is common to all VLTI/GRAVITY observations. It consists in extracting the complex visibilities from the raw data, using the ESO pipeline (Lapeyrere et al. 2014). The pipeline takes care of the background subtraction, flat-field correction, bad-pixel interpolation and P2VM multiplication (Tatulli et al. 2007). It also corrects the phase of the visibilities using the metrology data, and combines all DITs within each exposure. This last step performed by the pipeline (averaging of all DITs within each exposure file) is unwanted for exoplanet observations (see Section A.1). Thus, for the β Pic b observations, an intermediate file product generated by the pipeline is used: the "astrored" files, in which all DITs are kept separate. The complex visibilities contained in the "astrored" files are not corrected for the metrology and fringe-tracker zero-point, and thus the correction must be applied manually (see recipe in Nowak, 2019).

For each baseline b, and each time t (i.e. for each DIT), the wavelength-dependent complex visibility $\text{VISDATA}_{\text{onstar}}$ and $\text{VISDATA}_{\text{onplanet}}$ extracted by the pipeline (and with the above-mentionned corrections) are then "phase-referenced" to the star:

$$\underline{V}_{\text{onstar}} = |\text{VISDATA}_{\text{onstar}}| \tag{A.13}$$

$$\underline{V}_{\text{onplanet}} = \text{VISDATA}_{\text{onplanet}} \times e^{-i \arg (\text{VISDATA}_{\text{onstar}})} \tag{A.14}$$

where $\arg (\text{VISDATA}_{\text{onstar}})$ is the phase of the stellar complex visibility as measured by GRAVITY when the science fiber is positioned on the star.

This phase-referencing step is performed both on the star exposures, and on the planet exposures. When dealing with a star exposure, phase-referencing the visibility is mathematically equivalent to extracting the modulus of the visibility. But when dealing with on-planet exposures, a problem arises: the instrument does not simultaneously observe both the planet and the star. Thus, the quantity $\arg (\text{VISDATA}_{\text{onstar}})$ must be estimated from the available star exposures. In the observing strategy used for acquiring the β Pictoris data reported here, a star exposure was performed before and after each on-planet exposure. For each on-planet exposure, the phase reference is then simply estimated by taking the phase of the stellar complex visibily averaged on these two star exposures (before and after the on-planet observation).

A.3. A model for the on-planet visibility

In the absence of stellar flux, the on-planet visibility measured by the instrument and phase-referenced to the star can be written:

$$\underline{V}_{\text{onplanet}}(b, t, \lambda) = \underline{G}(b, t, \lambda)\underline{V}_{\text{planet}}(b, t, \lambda) \tag{A.15}$$

in which V_{planet} is the planet astrophysical visibility phase-referenced to the star, and G is the instrumental response. We note that the visibilities are not calibrated, meaning that the visibility at zero frequency is not 1, but the unnormalized flux. Therefore, as long as the planet remains unresolved by the instrument, its astrophysical visibility is given by:

$$\underline{V}_{\text{planet}}(b, t, \lambda) = S_{\text{planet}}(\lambda) \times e^{-i \frac{2\pi}{\lambda}(\Delta\alpha \times U + \Delta\delta \times V)} \tag{A.16}$$

in which (U, V) are the coordinates of baseline b in the UV-plane, $(\Delta\alpha, \Delta\delta)$ the sky-coordinates of the planet relative to the star, and $S_{\text{planet}}(\lambda)$ the spectrum of the planet.

Given the typical VLTI baseline lengths (between 45 and 130 m with the UTs), and the expected β Pic b planet-to-star separation at time of observation ($\simeq$ 140 mas), the exponential term in the above equation should produce significant oscillations of the complex visibility over the GRAVITY wavelength range (1.9 to 2.35 μm). But the phase-referenced on-planet visibilities extracted from our β Pic b observations show no such oscillations. The reason is that the data are dominated by remaining starlight, which needs to be taken into account.

To take into account the coherent starlight leaking into the fiber, Eq. (A.15) must be modified with an additional term, proportional to the stellar phase-referenced visibility V_{star}. In practice, since the leaking starlight does not originate in the direct coupling of the star to the fiber, but rather in the coupling of speckle noise to the fiber, this term needs to be multiplied by a polynomial in λ to account for its chromaticity. The model is now given by:

$$\begin{aligned}\underline{V}_{\text{onplanet}}(b, t, \lambda) &= \underline{Q}(b, t, \lambda)\underline{G}(b, t, \lambda)V_{\text{star}}(b, t, \lambda) \\ &+ \underline{G}(b, t, \lambda)\underline{V}_{\text{planet}}(b, t, \lambda)\end{aligned} \tag{A.17}$$

with $\lambda \to \underline{Q}(b, t, \lambda)$ a polynomial function in λ, whose coefficients vary with baseline b and time t.

If the star is not resolved by the instrument, its astrophysical phase-referenced visibility corresponds to its spectrum. If the star is partially resolved by the instrument, the spectrum needs to be multiplied by a term accounting for the resulting drop in visibility, which depends on the angular size of the star, limb-darkening model, etc. Explicitly separating these two terms, the referenced stellar astrophysical visibility can be written using the following equation, in which $S_\star(\lambda)$ is the star spectrum, and J a function accounting for the visibility drop due to the star geometry (typically, J is a bessel function of first order):

$$V_{\text{star}}(b, t, \lambda) = S_\star(\lambda)J(b, t, \lambda) \tag{A.18}$$

Going back to Eq. (A.17), the planet term in the right-hand side can be factored by $V_\star$ by introducing the planet-to-star contrast spectrum $C(\lambda) = S_{\text{planet}}(\lambda)/S_\star(\lambda)$:

$$\begin{aligned}\underline{V}_{\text{onplanet}} &= \underline{Q}\underline{G}V_{\text{star}} \\ &+ J^{-1}\underline{G}V_{\text{star}}C(\lambda)e^{-i\frac{2\pi}{\lambda}(\Delta\alpha U + \Delta\delta V)}\end{aligned} \tag{A.19}$$

The on-star equivalent of Eq. (A.17) is simpler, as the reference visibility observed on-star only depends on the stellar referenced visibility and the instrumental response:

$$\underline{V}_{\text{onstar}}(b, t, \lambda) = \underline{G}(b, t, \lambda)V_{\text{star}}(b, t, \lambda) \tag{A.20}$$

This provides a natural way to estimate the term $\underline{G}V_{\text{star}}$ in Eq. (A.19), and thus to calibrate $\underline{V}_{\text{onplanet}}$:

$$\begin{aligned}\underline{U}(b, t, \lambda) &= \frac{\underline{V}_{\text{onplanet}}}{\underline{V}_{\text{onstar}}}(b, t, \lambda) \\ &= \underline{Q} + J^{-1}C(\lambda)e^{-i\frac{2\pi}{\lambda}(\Delta\alpha U + \Delta\delta V)}\end{aligned} \tag{A.21}$$

Equation (A.21) shows how the physical quantities of interest (i.e. the contrast spectrum $C(\lambda)$ and the planet separation $\Delta\alpha, \Delta\delta$) are encoded in the on-planet data. Even

taking into account the filtering of the starlight by the off-axis fiber, as well as the only partly coherent nature of the speckle noise, the polynomial Q in the right-hand side of Eq. (A.21), which model the stellar residuals, is still a factor 20 to 30 superior to the planet signal. It is only because of the phase modulation naturally introduced by the planet separation that this planet signal can be retrieved. To do so, we proceed in two steps: we first extract the star-planet separation vector under some hypothesis on the contrast spectrum, and we then extract the contrast spectrum using the estimated separation vector. The two steps are iterated on time, to check the consistency of the results.

In matrix notations, the multiplications by $\underline{G}$, J^{-1}, the exponential, or even the polynomial Q can all be represented by diagonal-matrix multiplications. We write:

$$\underline{\mathbf{U}}_{b,t} = \sum_{k=0}^{m} \underline{a}_{b,t,k}\boldsymbol{\Lambda}^k\mathbb{1} + \mathbf{J}_{b,t}^{-1}\boldsymbol{\Phi}_{b,t}^{\Delta\alpha,\Delta\delta}\mathbf{C}, \qquad (A.22)$$

where m is the order of the polynomial Q, and the $\underline{a}_k$s are complex coefficients used to describe the polynomial. The vector $\mathbf{C}$ is defined from $C(\lambda)$ using the notations introduced in Section A.1, $\mathbb{1}$ is a column vector filled with 1's, and the matrices $\boldsymbol{\Lambda}$, $\mathbf{J}^{-1}$, and $\boldsymbol{\Phi}$ are all diagonal matrices of size $n_\lambda \times n_\lambda$ defined by:

$$\underline{\boldsymbol{\Lambda}} = \mathrm{diag}\left\{\lambda_1, \ldots, \lambda_{n_\lambda}\right\}$$
$$\mathbf{J}_{b,t}^{-1} = \mathrm{diag}\left\{J(b,t,\lambda_1)^{-1}, \ldots, J(b,t,\lambda_{n_\lambda})^{-1}\right\} \quad (A.23)$$
$$\boldsymbol{\Phi}_{b,t}^{\Delta\alpha,\Delta\delta} = \mathrm{diag}\left\{e^{-i\frac{2\pi}{\lambda_1}(\Delta\alpha U(b,t)+\Delta\delta V(b,t))}, \ldots\right\}$$

A.4. Extracting the astrometry

At the initial iteration, the planet to star contrast spectrum $\mathbf{C}$ in Eq. (A.22) can most generally be replaced by a flat spectrum. In the case of β Pictoris b, the temperature and surface gravity of the planet are known from previous work (Chilcote et al. 2017). Thus, the contrast spectrum $\mathbf{C}$ can be set to a model value. We use a BT-Settl model (Baraffe et al. 2015), at $T = 1700$ K and $\log(g/g_0) = 4.0$ (planet spectrum), divided by a BT-NextGen model (Hauschildt et al. 1999) at $T = 8000$ K and $\log(g/g_0) = 4.0$ (the star).

The contrast spectrum being set to a pre-determined value, Eq. (A.22) becomes a model at $(m+1) \times n_b n_{\mathrm{DIT}}$ complex parameters (the $\underline{a}_k$s), and 2 real parameters ($\Delta\alpha$ and $\Delta\delta$).

Interestingly, this model is linear in all the $\underline{a}_k$, and non-linear in $\Delta\alpha, \Delta\delta$. To fully benefit from this for the model-fitting, Eq. (A.22) can be re-arranged in a pseudo matrix form, with real parameters.
Introducing:

$$\mathbf{x}_{b,t} = \begin{pmatrix} \mathrm{Re}(\underline{a}_0) \\ \mathrm{Im}(\underline{a}_0) \\ \vdots \\ \mathrm{Re}(\underline{a}_m) \\ \mathrm{Im}(\underline{a}_m) \end{pmatrix} \qquad (A.24)$$

And $\underline{\mathbf{A}}_{b,t}^{\Delta\alpha,\Delta\delta}$ defined column by column:

$$\underline{\mathbf{A}}_{b,t}^{\Delta\alpha,\Delta\delta} = \left(\boldsymbol{\Lambda}^0\mathbb{1}, i\boldsymbol{\Lambda}^0\mathbb{1}, \ldots, \boldsymbol{\Lambda}^m\mathbb{1}, i\boldsymbol{\Lambda}^m\mathbb{1}, \underline{\boldsymbol{\Phi}}_{b,t}^{\Delta\alpha,\Delta\delta}\mathbf{C}\right) \quad (A.25)$$

We have:

$$\underline{\mathbf{U}}_{b,t} = \underline{\mathbf{A}}_{b,t}^{\Delta\alpha,\Delta\delta}\mathbf{x}_{b,t} \qquad (A.26)$$

For a given $\Delta\alpha$ and $\Delta\delta$, the corresponding best estimate of $\mathbf{x}_{b,t}$ (in the sense of the maximum likelihood) is given by:

$$\hat{\mathbf{x}}_{b,t} = \left(\mathrm{Re}\left(\underline{\mathbf{U}}_{b,t,2}^{\dagger}\underline{\mathbf{W}}_2^{-1}\underline{\mathbf{A}}_{b,t,2}\right)\left[\mathrm{Re}\left(\underline{\mathbf{A}}_{b,t,2}^{\dagger}\underline{\mathbf{W}}_2^{-1}\underline{\mathbf{A}}_{b,t,2}\right)\right]^{-1}\right)^T \tag{A.27}$$

where $\underline{\mathbf{W}}_2$ is a matrix composed of the covariance and pseudo covariance matrices of $\underline{\mathbf{U}}_{b,t}$ as defined in Eq. (A.11).

The log-likelihood $\log\mathcal{L}_{b,t}(\Delta\alpha,\Delta\beta)$, restricted to baseline b, DIT t, and to the nonlinear parameter $\Delta\alpha$ and $\Delta\delta$ is then given by the following equation, in which the dependance in $\Delta\alpha, \Delta\delta$ of the right-hand side is implicit in the definition of $\underline{\mathbf{A}}$ and $\hat{\mathbf{x}}$.

$$-\log\mathcal{L}_{b,t} = [\underline{\mathbf{U}}_{b,t} - \underline{\mathbf{A}}_{b,t}\hat{\mathbf{x}}_{b,t}]_2^{\dagger}\underline{\mathbf{W}}_{b,t,2}^{-1}[\underline{\mathbf{U}}_{b,t} - \underline{\mathbf{A}}_{b,t}\hat{\mathbf{x}}_{b,t}]_2 \tag{A.28}$$

The total log-likelihood can be obtained by summing over all baselines b and DITs t:

$$-\log\mathcal{L}(\Delta\alpha,\Delta\delta) = -\sum_{b,t}\log\mathcal{L}_{b,t}(\Delta\alpha,\Delta\delta) \qquad (A.29)$$

With the expression of $\underline{\mathbf{A}}_{b,t}$ from Eq. (A.25) and $\hat{\mathbf{x}}$ from Eq. (A.27), this gives a closed-form expression of the log-likelihood in $\Delta\alpha$, $\Delta\delta$, from which a map can be calculated, in order to extract the best estimate with the associated error bars.

A.5. Extracting the spectrum

Extracting the spectrum from the GRAVITY observations is more difficult than extracting the astrometry for two reasons: first, due to the dimensionality of the problem ($n_\lambda > 200$), a $\log\mathcal{L}$ map approach is impractical; second, the stellar residuals affecting the on-planet visibility can lead to a degenerated solution for the contrast spectrum $\mathbf{C}$.

The impact of the stellar residuals on the contrast spectrum can be quantified by using Eq. (A.22) again. The calculations in the rest of this section are simpler when considering all visibilities shifted to the planet position. To do so, we multiply both sides of Eq. (A.22) by the inverse of $\underline{\boldsymbol{\Phi}}_{b,t}^{\Delta\alpha,\Delta\delta}$. From now on, we will denote $\tilde{\underline{\mathbf{V}}}_{b,t}$ any "shifted" visibility $\underline{\mathbf{V}}_{b,t}$. Eq. (A.22) becomes:

$$\tilde{\underline{\mathbf{U}}}_{b,t} = \sum_{k=0}^{m}\underline{a}_{b,t,k}\boldsymbol{\Lambda}^k\tilde{\mathbb{1}} + \mathbf{J}_{b,t}^{-1}\mathbf{C} \qquad (A.30)$$

From there, we can introduce the subspace $\mathbb{C}^m[\boldsymbol{\Lambda}]\tilde{\mathbb{1}}$ of $\mathbb{C}^{n_\lambda}$, defined as the subspace generated by the family of $m+1$ vectors $\boldsymbol{\Lambda}^0\tilde{\mathbb{1}}, \ldots, \boldsymbol{\Lambda}^m\tilde{\mathbb{1}}$ (i.e., the subspace of vectors which are linear combinations of these $m+1$ vectors). Introducing this subspace if of course motivated by the fact that the stellar residual term is part of it. We can then introduce the projector matrix orthogonal to this subspace, which we denote[3] $\underline{\mathbf{P}}_{\mathbb{C}^m[\boldsymbol{\Lambda}]\tilde{\mathbb{1}}_{b,t}}$. Projecting Eq. (A.22) then gives:

$$\underline{\mathbf{P}}_{\mathbb{C}^m[\boldsymbol{\Lambda}]\tilde{\mathbb{1}}_{b,t}}\tilde{\underline{\mathbf{U}}}_{b,t} = \underline{\mathbf{P}}_{\mathbb{C}^m[\boldsymbol{\Lambda}]\tilde{\mathbb{1}}_{b,t}}\mathbf{J}_{b,t}^{-1}\mathbf{C} \qquad (A.31)$$

---

[3] The vector $\tilde{\mathbb{1}}$ hides a dependancy in b,t through the matrix $\boldsymbol{\Phi}_{b,t}^{\Delta\alpha,\Delta\delta}$ used to define the tilded vector. As this dependency is important in the following calculation, it is made explicit by using $\tilde{\mathbb{1}}_{b,t}$ instead

The new Eq. (A.31) is a representation of the exact information content of Eq. (A.22) regarding the contrast spectrum. Since $\underline{\mathbf{P}}_{\mathbb{C}^m[\Lambda]\tilde{\underline{\mathbb{1}}}}$ is a projector matrix, it is necessarily of rank $< n_\lambda$, with an exact value which depends on the dimension of the subspace generated by the $\boldsymbol{\Lambda}^k \tilde{\underline{\mathbb{1}}}$. Thus, Eq. (A.31) is not invertible, and the contrast spectrum cannot be fully recovered from it.

Fortunately, the dataset acquired on β Pictoris b contains several baselines and DITs. For each baseline and each DIT, the projector matrix is different, since it depends on the matrix $\boldsymbol{\Phi}_{b,t}$ through the vector $\tilde{\underline{\mathbb{1}}}_{b,t}$. These variations can be leveraged to unambiguously recover the complete contrast spectrum $\mathbf{C}$.

The proper way to proceed is to start by extracting a set of linearly independant equations from Eq. (A.31). This can be done by using a diagonal representation of the projector matrix, for example using a singular value decomposition. We can introduce an hermitian matrix $\underline{\mathbf{H}}_{b,t}$ and a diagonal matrix $\mathbf{D}_{b,t}$ such that:

$$\underline{\mathbf{H}}_{b,t}\underline{\mathbf{P}}_{\mathbb{C}^m[\Lambda]\tilde{\underline{\mathbb{1}}}_{b,t}}\underline{\mathbf{H}}_{b,t}^\dagger = \mathbf{D}_{b,t} \tag{A.32}$$

$$\underline{\mathbf{H}}_{b,t}\underline{\mathbf{H}}_{b,t}^\dagger = \mathbf{I} \tag{A.33}$$

Since $\underline{\mathbf{P}}_{\mathbb{C}^m[\Lambda]\tilde{\underline{\mathbb{1}}}_{b,t}}$ is a projector matrix, its eigenvalues are either 1 or 0. We can assume $\mathbf{D}$ to be of the form:

$$\mathbf{D}_{b,t} = \begin{pmatrix} \mathbf{I}_{r(b,t)} & 0 \\ 0 & 0 \end{pmatrix} \tag{A.34}$$

where $\mathbf{I}_r(b,t)$ is the identity matrix of size given by the rank of the projector: $r(b,t) = \mathrm{rank}(\underline{\mathbf{P}}_{\mathbb{C}^m[\Lambda]\tilde{\underline{\mathbb{1}}}_{b,t}})$.

The matrix $\underline{\mathbf{H}}_{b,t}$ can also be written in blocks:

$$\underline{\mathbf{H}}_{b,t} = \begin{pmatrix} \underline{\mathbf{H}}_{b,t}^{11} & \underline{\mathbf{H}}_{b,t}^{12} \\ \underline{\mathbf{H}}_{b,t}^{21} & \underline{\mathbf{H}}_{b,t}^{22} \end{pmatrix} \tag{A.35}$$

From there, multiplying both sides of Eq. (A.31) by $\underline{\mathbf{H}}_{b,t}$, noting that $\underline{\mathbf{H}}_{b,t}\underline{\mathbf{P}}_{\mathbb{C}^m[\Lambda]\tilde{\underline{\mathbb{1}}}_{b,t}} = \mathbf{D}_{b,t}\underline{\mathbf{H}}_{b,t}$, and using a block calculation gives:

$$\begin{bmatrix} \underline{\mathbf{H}}_{b,t}^{11} & \underline{\mathbf{H}}_{b,t}^{12} \end{bmatrix} \tilde{\underline{\mathbf{U}}}_{b,t} = \begin{bmatrix} \underline{\mathbf{H}}_{b,t}^{11} & \underline{\mathbf{H}}_{b,t}^{12} \end{bmatrix} \mathbf{J}_{b,t}\mathbf{C} \tag{A.36}$$

and a dummy equation $0 = 0$.

Since $r(b,t) < n_\lambda$, the linearly independent system defined by Eq. (A.36) is underdetermined (the matrix $\begin{bmatrix} \underline{\mathbf{H}}_{b,t}^{11} & \underline{\mathbf{H}}_{b,t}^{12} \end{bmatrix}$ has more columns than rows).

To invert the problem, it is necessary to combine all the equations obtained for the different baselines b and t. In matrix notation, this is just a matter of concatenating all the sub-matrices:

$$\underline{\mathbf{H}} = \begin{pmatrix} \begin{bmatrix} \mathbf{H}_{1,1}^{11} & \mathbf{H}_{1,1}^{12} \end{bmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{bmatrix} \mathbf{H}_{n_b,n_{\mathrm{DIT}}}^{11} & \mathbf{H}_{n_b,n_{\mathrm{DIT}}}^{12} \end{bmatrix} \end{pmatrix} \tag{A.37}$$

$$\tilde{\underline{\mathbb{U}}} = \begin{pmatrix} \tilde{\underline{\mathbf{U}}}_{1,1} \\ \tilde{\underline{\mathbf{U}}}_{2,1} \\ \vdots \\ \tilde{\underline{\mathbf{U}}}_{n_b,n_\lambda} \end{pmatrix} \tag{A.38}$$

$$\mathbb{J} = \begin{pmatrix} \mathbf{J}_{1,1}^{-1} \\ \mathbf{J}_{2,1}^{-1} \\ \vdots \\ \mathbf{J}_{n_b,n_\lambda}^{-1} \end{pmatrix} \tag{A.39}$$

This gives:

$$\underline{\mathbf{H}}\tilde{\underline{\mathbb{U}}} = \underline{\mathbf{H}}\mathbb{J}\mathbf{C} \tag{A.40}$$

Which has the form of a linear problem:

$$\mathbb{Y} = \mathbb{H}\mathbf{C} \tag{A.41}$$

Where $\mathbb{Y}$ is a linear transformation of the calibrated visibilities defied by:

$$\mathbb{Y} = \underline{\mathbf{H}}\tilde{\underline{\mathbb{U}}} \tag{A.42}$$

And $\mathbb{H}$ is the collapsed matrix:

$$\mathbb{H} = \begin{pmatrix} \begin{bmatrix} \mathbf{H}_{1,1}^{11} & \mathbf{H}_{1,1}^{12} \end{bmatrix} \mathbf{J}_{1,1}^{-1} \\ \begin{bmatrix} \mathbf{H}_{2,1}^{11} & \mathbf{H}_{2,1}^{12} \end{bmatrix} \mathbf{J}_{2,1}^{-1} \\ \vdots \\ \begin{bmatrix} \mathbf{H}_{n_b,n_{\mathrm{DIT}}}^{11} & \mathbf{H}_{n_b,n_{\mathrm{DIT}}}^{12} \end{bmatrix} \mathbf{J}_{n_b,n_{\mathrm{DIT}}}^{-1} \end{pmatrix} \tag{A.43}$$

The problem defined by Eq. (A.41) can be solved using the maximum likelihood formalism, adapted to complex random variables. The expression of the maximum likelihood solution is given by Eq. (A.9). For the contrast spectrum, we have:

$$\hat{\mathbf{C}} = \left(\mathrm{Re}\left(\tilde{\mathbb{Y}}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) \left[\mathrm{Re}\left(\mathbb{H}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) \right]^{-1} \right)^T \tag{A.44}$$

where the extended vectors and matrices $\tilde{\mathbf{Y}}_2$, $\mathbb{H}_2$, and $\underline{\mathbf{W}}_2$, are defined by:

$$\tilde{\mathbf{Y}}_2 = \begin{pmatrix} \tilde{\mathbf{Y}} \\ \tilde{\mathbf{Y}}^* \end{pmatrix} \quad \mathbb{H}_2 = \begin{pmatrix} \mathbb{H} \\ \mathbb{H}^* \end{pmatrix} \quad \underline{\mathbf{W}}_2 = \begin{pmatrix} \mathbf{W} & \mathbf{Z} \\ \mathbf{Z}^\dagger & \mathbf{W}^* \end{pmatrix} \tag{A.45}$$

The uncertainty on this best estimate of the contrast spectrum can be obtained with a direct error propagation all the way to the real covariance matrix on $\hat{\mathbf{C}}$. The covariance and pseudo covariance matrices of $\mathbb{Y}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2$ are given by:

$$\mathrm{cov}\left(\mathbb{Y}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) = \mathbb{H}_2^\dagger \underline{\mathbf{W}}_2^{-1\dagger} \mathrm{cov}(\mathbb{Y}_2) \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \tag{A.46}$$

$$\mathrm{pcov}\left(\mathbb{Y}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) = \mathbb{H}_2^T \underline{\mathbf{W}}_2^{-1T} \mathrm{pcov}(\mathbb{Y}_2) \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \tag{A.47}$$

A proper combination of these covariance and pseudo-covariance matrices gives the covariance matrix of the real part of $\mathbb{Y}_2 \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2$:

$$\begin{aligned} &\mathrm{cov}\left\{ \mathrm{Re}\left(\mathbb{Y}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) \right\} = \\ &\frac{1}{2}\mathrm{Re}\left(\mathrm{cov}\left(\mathbb{Y}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) + \mathrm{pcov}\left(\mathbb{Y}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) \right) \end{aligned} \tag{A.48}$$

Which can then be propagated to give the final full covariance matrix on $\hat{\mathbf{C}}$:

$$\begin{aligned} &\mathrm{cov}\left(\hat{\mathbf{C}} \right) = \\ &\mathrm{Re}\left(\mathbb{H}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right)^{-1T} \mathrm{cov}\left(\mathrm{Re}\left(\tilde{\mathbb{Y}}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right) \right) \mathrm{Re}\left(\mathbb{H}_2^\dagger \underline{\mathbf{W}}_2^{-1} \mathbb{H}_2 \right)^{-1} \end{aligned} \tag{A.49}$$
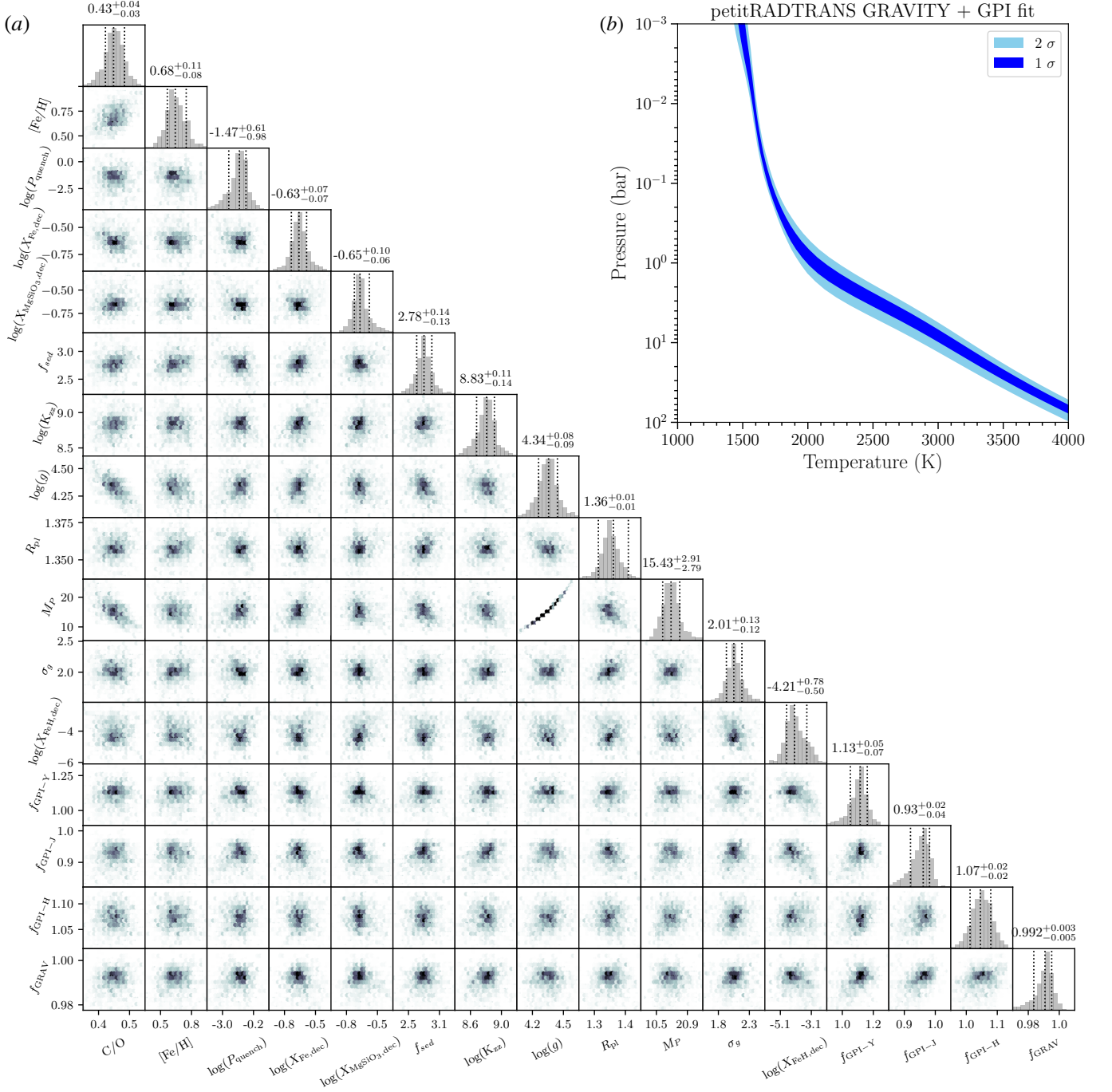
Fig. B.1. *Panel (a)*: projected 2-d posterior of the GRAVITY+GPI fit with petitRADTRANS (spectrum shown in Figure 5), described in Section 4.3. See the text in Section B for a description of the parameters. *Panel (b)*: pressure-temperature envelopes obtained for the same retrieval. At every pressure, we plot the 16 to 84-percentile envelopes in dark blue, and the 2.5 to 97.5 percentile envelopes in light blue. If the temperature values were following a Gauss distribution, this would correspond to the 1 and 2 σ envelopes, respectively.

Appendix B: Posterior of the petitRADTRANS fit

In Figure B.1 we show the posteriors of the GRAVITY+GPI fit with petitRADTRANS, described in Section 4.3. Panel (a) shows the corner plot for all but the temperature nuisance parameters. Panel (b) shows the retrieved temperature uncertainty envelopes.

The parameters shown are the following: the C/O, adjusted by changing the oxygen abundance at a given [Fe/H].

The metallicity [Fe/H], which was used to scale the number fraction of all atomic elements (except H and He) by $10^{[Fe/H]}$. The quench pressure P_{quench} of the atmosphere, here converged to a low enough value such that non-equilibrium chemistry effects are negligible. The mass fractions of Fe and $MgSiO_3$ at the cloud base. Here they are expressed in units of a log-ed decrease factor, which gets multiplied with the maximally allowed mass fraction, based on the elemental composition of the atmosphere. The

19

stoichiometric factors of $MgSiO_3$ are used for finding this upper limit. The cloud settling parameter f_{sed}, as described in Section 4.3. The log-ed eddy diffusion coefficient K_{zz}, in units of $cm^2 \, s^{-1}$. This is used for calculating the cloud particle size, as described in Section 4.3. The planet's surface gravity $\log(g)$. The planetary radius R_P, in units of Jupiter radii. The planetary mass M_P, in units of Jupiter masses. This is calculated using the sampled $\log(g)$ and R_P values. The width of the log-normal cloud particle size distribution σ_g, as described in Section 4.3. The vertically constant mass fraction of FeH, expressed in units of a log-ed decrease factor. This factor gets multiplied with the maximally allowed mass fraction of FeH, based on the elemental composition of the atmosphere and the Fe atoms not yet incorporated into the Fe clouds. Finally the f_{GPI-Y}, f_{GPI-J}, f_{GPI-H} and f_{GRAV} factors describe the multiplicative scaling of the individual bands, which were allowed to vary by 50 % in the case of GPI, and by 5 % in the case of GRAVITY.

The temperature envelopes in Panel (b) of Figure B.1 are obtained by plotting, at every pressure, the 16 to 84-percentile envelopes, and the 2.5 to 97.5 percentile envelopes. If the temperature values at a given pressure layer were following a normal distribution, this would correspond to the 1 and 2 σ envelopes, respectively.

References

Ackerman, A. S. & Marley, M. S. 2001, ApJ, 556, 872
Ali-Dib, M., Mousis, O., Petit, J.-M., & Lunine, J. I. 2014, ApJ, 785, 125
Alibert, Y., Mousis, O., Mordasini, C., & Benz, W. 2005, ApJ, 626, L57
Allard, F., Hauschildt, P. H., Alexander, D. R., Tamanai, A., & Schweitzer, A. 2001, ApJ, 556, 357
Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765
Andrews, S. M. & Williams, J. P. 2005, ApJ, 619, L175
Ansdell, M., Williams, J. P., Manara, C. F., et al. 2017, AJ, 153, 240
Asplund, M. 2005, ARA&A, 43, 481
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701
Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42
Baudino, J.-L., Bézard, B., Boccaletti, A., et al. 2015, A&A, 582, A83
Baudino, J.-L., Mollière, P., Venot, O., et al. 2017, ApJ, 850, 150
Béjar, V. J. S. & Martín, E. L. 2018, Brown Dwarfs and Free-Floating Planets in Young Stellar Clusters, 92
Bell, C. P. M., Mamajek, E. E., & Naylor, T. 2015, MNRAS, 454, 593
Blunt, S., Wang, J., Angelo, I., et al. 2019, arXiv e-prints, arXiv:1910.01756
Bodenheimer, P. 1974, Icarus, 23, 319
Bodenheimer, P., Grossman, A. S., Decampli, W. M., Marcy, G., & Pollack, J. B. 1980, Icarus, 41, 293
Bonnefoy, M., Boccaletti, A., Lagrange, A.-M., et al. 2013, A&A, 555, A107
Bonnefoy, M., Marleau, G.-D., Galicher, R., et al. 2014, A&A, 567, L9
Brandt, T. D. 2018, ApJS, 239, 31
Burningham, B., Marley, M. S., Line, M. R., et al. 2017, MNRAS, 470, 1177
Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, ApJ, 542, 464
Charnay, B., Bézard, B., Baudino, J.-L., et al. 2018, ApJ, 854, 172
Chauvin, G., Lagrange, A. M., Beust, H., et al. 2012, A&A, 542, A41
Chilcote, J., Barman, T., Fitzgerald, M. P., et al. 2015, ApJ, 798, L3
Chilcote, J., Pueyo, L., De Rosa, R. J., et al. 2017, AJ, 153, 182
Cridland, A. J., Pudritz, R. E., & Alessi, M. 2016, MNRAS, 461, 3274
Decampli, W. M. & Cameron, A. G. W. 1979, Icarus, 38, 367
Dupuy, T. J., Brandt, T. D., Kratter, K. M., & Bowler, B. P. 2019, ApJ, 871, L4

Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2016, A&A, 595, A83
Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2018, A&A, 613, A14
Feroz, F. & Hobson, M. P. 2008, MNRAS, 384, 449
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gravity Collaboration, Abuter, R., Accardo, M., et al. 2017, A&A, 602, A94
Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018, A&A, 618, L10
Gravity Collaboration, Lacour, S., Nowak, M., et al. 2019, A&A, 623, L11
Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, AJ, 132, 161
Hansen, J. E. 1971, Journal of Atmospheric Sciences, 28, 1400
Hauschildt, P. H., Allard, F., & Baron, E. 1999, ApJ, 512, 377
Helled, R., Podolak, M., & Kovetz, A. 2006, Icarus, 185, 64
Helled, R. & Schubert, G. 2008, Icarus, 198, 156
Helled, R. & Schubert, G. 2009, ApJ, 697, 1256
Helling, C., Dehn, M., Woitke, P., & Hauschildt, P. H. 2008, ApJ, 675, L105
Helling, C. & Woitke, P. 2006, A&A, 455, 325
Helling, C., Woitke, P., Rimmer, P. B., et al. 2014, Life, 4, 142
Holweger, H., Hempel, M., van Thiel, T., & Kaufer, A. 1997, A&A, 320, L49
Kervella, P., Arenou, F., Mignard, F., & Thévenin, F. 2019, A&A, 623, A72
Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, Science, 339, 1398
Kreidberg, L., Line, M. R., Bean, J. L., et al. 2015, ApJ, 814, 66
Lacour, S., Dembet, R., Abuter, R., et al. 2019, A&A, 624, A99
Lagrange, A.-M., Boccaletti, A., Langlois, M., et al. 2019a, A&A, 621, L8
Lagrange, A. M., De Bondt, K., Meunier, N., et al. 2012, A&A, 542, A18
Lagrange, A.-M., Keppler, M., Meunier, N., et al. 2018, A&A, 612, A108
Lagrange, A. M., Meunier, N., Rubini, P., et al. 2019b, Nature Astronomy, 421
Lanz, T., Heap, S. R., & Hubeny, I. 1995, ApJ, 447, L41
Lapeyrere, V., Kervella, P., Lacour, S., et al. 2014, in Proc. SPIE, Vol. 9146, Optical and Infrared Interferometry IV, 91462D
Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, AJ, 154, 91
Line, M. R., Marley, M. S., Liu, M. C., et al. 2017, ApJ, 848, 83
Line, M. R., Teske, J., Burningham, B., Fortney, J. J., & Marley, M. S. 2015, ApJ, 807, 183
Lissauer, J. J. & Stevenson, D. J. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil, 591–606
Madhusudhan, N., Amin, M. A., & Kennedy, G. M. 2014, ApJ, 794, L12
Madhusudhan, N., Bitsch, B., Johansen, A., & Eriksson, L. 2017, MNRAS, 469, 4102
Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011, Nature, 469, 64
Madhusudhan, N. & Seager, S. 2009, ApJ, 707, 24
Marboeuf, U., Thiabaud, A., Alibert, Y., Cabral, N., & Benz, W. 2014a, A&A, 570, A36
Marboeuf, U., Thiabaud, A., Alibert, Y., Cabral, N., & Benz, W. 2014b, A&A, 570, A35
Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007, ApJ, 655, 541
Marley, M. S., Saumon, D., Cushing, M., et al. 2012, ApJ, 754, 135
Millar-Blanchaer, M. A., Graham, J. R., Pueyo, L., et al. 2015, ApJ, 811, 18
Mollière, P. & Mordasini, C. 2012, A&A, 547, A105
Mollière, P. & Snellen, I. A. G. 2019, A&A, 622, A139
Mollière, P., van Boekel, R., Bouwman, J., et al. 2017, A&A, 600, A10
Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, ApJ, 813, 47
Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, A&A, 627, A67
Mordasini, C., Marleau, G.-D., & Mollière, P. 2017, A&A, 608, A72
Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, ApJ, 832, 41
Morley, C. V., Marley, M. S., Fortney, J. J., et al. 2014, ApJ, 787, 78
Morzinski, K. M., Males, J. R., Skemer, A. J., et al. 2015, ApJ, 815, 108

Nielsen, E. L., Liu, M. C., Wahhaj, Z., et al. 2014, ApJ, 794, 158
Notsu, S., Akiyama, E., Booth, A., et al. 2019, ApJ, 875, 96
Nowak, M. 2019, PhD thesis, Universit PSL
Öberg, K. I. & Bergin, E. A. 2016, ApJ, 831, L19
Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJ, 743, L16
Owen, T., Mahaffy, P., Niemann, H. B., et al. 1999, Nature, 402, 269
Pascucci, I., Testi, L., Herczeg, G. J., et al. 2016, ApJ, 831, 125
Qi, C., Öberg, K. I., Andrews, S. M., et al. 2015, ApJ, 813, 128
Samland, M., Mollière, P., Bonnefoy, M., et al. 2017, A&A, 603, A57
Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, Nature, 509, 63
Snellen, I. A. G. & Brown, A. G. A. 2018, Nature Astronomy, 2, 883
Spiegel, D. S., Burrows, A., & Milsom, J. A. 2011, ApJ, 727, 57
Tatulli, E., Millour, F., Chelli, A., et al. 2007, A&A, 464, 29
Thébault, P. & Beust, H. 2001, A&A, 376, 621
Thiabaud, A., Marboeuf, U., Alibert, Y., et al. 2014, A&A, 562, A27
van der Bliek, N. S., Manfroid, J., & Bouchet, P. 1996, A&AS, 119, 547
van Leeuwen, F. 2007, A&A, 474, 653
Vousden, W. D., Farr, W. M., & Mandel, I. 2016, MNRAS, 455, 1919
Walsh, C., Nomura, H., & van Dishoeck, E. 2015, A&A, 582, A88
Wang, J. J., Graham, J. R., Dawson, R., et al. 2018, AJ, 156, 192
Wang, J. J., Graham, J. R., Pueyo, L., et al. 2016, AJ, 152, 97
Woitke, P. & Helling, C. 2003, A&A, 399, 297
Zahnle, K. J. & Marley, M. S. 2014, ApJ, 797, 41
Zalesky, J. A., Line, M. R., Schneider, A. C., & Patience, J. 2019, ApJ, 877, 24
Zieba, S., Zwintz, K., Kenworthy, M. A., & Kennedy, G. M. 2019, A&A, 625, L13

[1] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cité, 5 place Jules Janssen, 92195 Meudon, France

[2] European Southern Observatory, Karl-Schwarzschild-Strae 2, 85748 Garching, Germany

[3] Max Planck Institute for extraterrestrial Physics, Giessenbachstrae 1, 85748 Garching, Germany

[4] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany

[5] Max Planck Institute for Radio Astronomy, Auf dem Hügel 69, 53121 Bonn, Germany

[6] 1st Institute of Physics, University of Cologne, Zülpicher Strae 77, 50937 Cologne, Germany

[7] Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France

[8] Universidade de Lisboa - Faculdade de Ciências, Campo Grande, 1749-016 Lisboa, Portugal

[9] Sterrewacht Leiden, Leiden University, Postbus 9513, 2300 RA Leiden, The Netherlands

[10] Departments of Physics and Astronomy, Le Conte Hall, University of California, Berkeley, CA 94720, USA

[11] CENTRA - Centro de Astrofísica e Gravitação, IST, Universidade de Lisboa, 1049-001 Lisboa, Portugal

[12] Space Telescope Science Institute, Baltimore, MD 21218, USA

[13] STAR Institute, Université de Liège, Allée du Six Août 19c, B-4000 Liège, Belgium

[14] Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

[15] Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA, United Kingdom