

# Upper limits on CN from exocomets transiting $\beta$ Pictoris

M.A. Kenworthy<sup>1</sup>, E. de Mooij<sup>2</sup>, A. Brandeker<sup>3</sup>, C. Opitom<sup>4</sup>, F. Kiefer<sup>5</sup>, and A. Fitzsimmons<sup>2</sup>

<sup>1</sup> Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands  
e-mail: kenworthy@strw.leidenuniv.nl

<sup>2</sup> Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, BT7 1NN Belfast, UK

<sup>3</sup> Institutionen för astronomi, Stockholms universitet, AlbaNova universitetscentrum, 106 91, Stockholm, Sweden

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

<sup>5</sup> LESIA, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92190 Meudon, France

Received 7 Feb 2025; Accepted 13 Apr 2025

## ABSTRACT

**Context.** The young (23 Myr) nearby (19.4 pc) star  $\beta$  Pictoris hosts an edge-on debris disk with two gas giant exoplanets in orbit around it. Many transient absorption features have been detected in the rotationally broadened stellar lines, which are thought to be the coma of infalling exocomets crossing the line of sight towards Earth.

**Aims.** In the Solar System, the molecule cyanoogen (CN) and its associated ionic species are one of the most detectable molecules in the coma and tails of comets. We perform a search for cyanogen in the spectra of  $\beta$  Pic to detect or put an upper limit on this molecule's presence in a young, highly active planetary system.

**Methods.** We divide twenty year's worth of High Accuracy Radial Velocity Planet Searcher (HARPS) spectra into those with strong exocomet absorption features, and those with only stellar lines. The high signal-to-noise stellar spectrum normalises out the stellar lines in the exocomet spectra, which are then shifted and stacked on the deepest exocomet absorption features to produce a high signal-to-noise exocomet spectrum, and search for the CN band head using a model temperature dependent cross-correlation template.

**Results.** We do not detect CN in our data, and place a temperature and broadening dependent  $5\sigma$  upper limit between  $10^{12}$  cm<sup>-2</sup> and  $10^{13}$  cm<sup>-2</sup>, to be compared to the typical  $10^9$  -  $10^{10}$  cm<sup>-2</sup> expected from scaling of the values in the Solar System comets.

**Conclusions.**

**Key words.** comets — transits, spectroscopy, stars: individual: Beta Pictoris

## 1. Introduction

The formation and evolution of planets and their attendant moons is thought to occur in the first several million years of a planetary system. Within our Solar System, comets deliver volatiles from the outermost reaches of the Oort cloud (O'Brien et al. 2018), past the gas giants and terrestrial planets, even as close as a few solar radii. It is therefore important to understand the distribution, nature and chemical composition of exocomets in other planet forming systems.

One of the closest young planetary systems is  $\beta$  Pictoris, which has been studied intensively both with direct imaging and with spectroscopy.  $\beta$  Pic is a young ( $\sim 25$  Myr; see Lee et al. 2024, for discussions), nearby ( $19.63 \pm 0.06$  pc, Lindegren et al. 2021) and early type star (A6V; Gray et al. 2006) which hosts an edge-on debris disk and at least two exoplanets, making it one of the most intensely studied young planetary systems. The star is a rapid rotator ( $124 \pm 3$  km s<sup>-1</sup>; Koen et al. 2003) seen close to equator on. Amongst these, the Ca H and K lines show deep absorption down to 0.1 of the continuum level. Deep within this rotationally broadened line there is a consistent and stable narrow absorption feature with a full width half maximum (FWHM) of a few km s<sup>-1</sup>, attributed to the presence of the circumstellar material (Hobbs et al. 1985; Vidal-Madjar et al. 1986).

Transient absorption features appearing in the rotationally broadened Ca H and K features were proposed to be due to the transit of evaporating exocomets in front of the star (Ferlet et al. 1987). These features appear and disappear on the timescale of

hours to days, with some features lasting up to tens of days and have red-shifted velocities from eighty kilometres per second, all the way down to (and merging with) the circumstellar line, with a few appearing on the blueward side (Lagrange-Henri et al. 1992) with widths from a few up to tens of kilometres per second. The absorption can be described by the transit of evaporating bodies that pass within a few stellar radii of the star, resulting in sublimation of material that produce a large coma containing calcium ions which are then seen as an additional absorption component (Beust et al. 1991). Intensive spectroscopic monitoring over several hours show both multiple components transiting the disk of the star (Beust et al. 1996) and changes of radial velocity of several components are consistent with Keplerian elliptical orbits (Kennedy 2018) implying at least two (Kiefer et al. 2014; Pollard 2024) or more (Heller 2024) distinct kinematic families of Beta Pictoris exocomets. At least one event is derived with the orbit parallel to the rotational axis of the star (Tobin et al. 2019).

From the multiple detections of exocomets in spectroscopy towards  $\beta$  Pic, the broadband detection of exocomet tails transiting a star were predicted in Lecavelier Des Etangs et al. (1999). A chain of seven exocomets were proposed to explain the shape of one eclipse seen towards KIC 8462852 in Kiefer et al. (2017), and single exocomet transits were subsequently detected towards the star KIC 3542116 (Rappaport et al. 2018) with the Kepler satellite (Borucki et al. 2010) and in TESS data towards  $\beta$  Pic (Zieba et al. 2019). Observations of  $\beta$  Pic in subsequent TESS sectors revealed over two dozen transits whose depth follows a power law distribution similar to those of the Solar System

(Lecavelier des Etangs et al. 2022; Pavlenko et al. 2022). Automated searches in photometry (Kennedy et al. 2019) and spectroscopy (Bendahane-West et al. 2024) for additional exocomet systems have been met with limited success.

To the contrary of what is observed in the case of Beta Picoris exocomets, most molecular and atomic signatures detected in the coma of Solar System comets are emission lines (or bands) instead of absorptions. Spectroscopic observations of Solar System comets when they are very close to the Sun are rare due to observational difficulties. The only comet spectroscopically observed closer than 0.2 au from the Sun is C/1965 S1 Ikeya-Seki. For this comet, it was noticed that the main emission bands or lines detected in its optical spectrum dramatically changed while the comets approached the Sun. At distances greater than 0.4 au, CH, CN, C<sub>2</sub>, and C<sub>3</sub> emission bands dominated the spectrum. However, at 0.15 au from the Sun, the only molecular emission detected was CN, while numerous atomic lines were visible ([O] I, Na I, K I, Ca II, Cr I, Mn I, Fe I, Ni I, Cu I, V I; Dufay et al. 1965; Thackeray et al. 1966; Preston 1967; Slaughter 1969). Additionally, a tail of Fe I atoms was proposed to explain a linear feature imaged in the tail of sungrazing comet C/2016 P1 McNaught 2006P1 (Fulle et al. 2007) and Na and K were detected for comet C/2011 L4 (Fulle et al. 2013) at 0.46 au.

On the other hand, Na I D-line emission has been observed in many comets within  $r \sim 1$  au of the Sun, where either the emission was much brighter than the foreground sky emission, or the comet was observed at high spectral resolution enabling separation from the telluric lines (Cremonese et al. 2002; Schmidt 2016). The high efficiency of the Na I D transition makes it easily detectable, even for relatively low column density. A notable discovery was a spectacular tail of Na I associated with comet C/1995 O1 Hale-Bopp at  $r \sim 1$  au, formed by radiation pressure on the neutral atoms before ionisation (Cremonese et al. 1997). Recently, Hui (2023) also demonstrated that the activity of near-Earth asteroid Phaeton, with a perihelion of only 0.14 au, is associated with emission of sodium.

The detection of these atoms only at small heliocentric distances imply that Ca II and Fe I are predominantly locked up within cometary dust grains, and only observed when a comet is close enough to the Sun to allow their sublimation. This occurs at  $r \leq 0.03 - 0.06$  au for a dust grain with Bond albedo  $\sim 0.01$  and a sublimation temperature of  $T \approx 1600$  K, depending on the thermal model assumed. On the other hand, the origin of the Na I tail observed in numerous comets at larger distances from the Sun is still not very well understood (Cremonese et al. 2002). Release from dust grains as well as molecular processes have been suggested as possible sources of Na I in the coma and tail of comets and explain its presence at large solar distances (Cremonese et al. 1997).

For the majority of Solar System comets, their optical spectra are dominated by the reflected sunlight from dust grains in their comae, plus molecular and atomic emission lines (e.g. Hyland et al. 2019). Most of the neutral and ionic molecular emission results from resonance fluorescence with solar photons. Some species, such as H<sub>2</sub>O<sup>+</sup> and CO<sup>+</sup> result from ionization of parent molecules (i.e. molecules released directly by the sublimation of nuclear ices), all other molecules observed at optical wavelengths are second generation species, often created via photodissociation of parent molecules or released by organic-rich grains in the coma. Another process also observed in optical spectrum of comets is prompt atomic emission ([O] I, [C] I, and [N] I). Those forbidden transitions are the result of atoms produced in an excited state by the photodissociation of parent molecules in the coma.

Of all cometary gas emission, the highest signal to noise normally occurs with the CN (0-0) band at  $\lambda \approx 3880$  Å. It is usually the first emission detected in the optical spectrum of a comet as it approaches the Sun, and as been detected in the comae of comets observed at relatively large distances from the Sun (Bus et al. 1991; Fitzsimmons & Cartwright 1996). The relative abundance in the coma of most comets of OH/CN  $\approx 500$  (A'Hearn et al. 1995), together with its high fluorescence scattering efficiency factor of approximately 2.6 photons per second per molecule at 1 au from the Sun and  $\dot{r} \approx 0$  (Schleicher 2010), gives rise to a clear signature of a cometary gas coma. Historically, the CN was believed to be created from photodissociation of HCN, which is present in abundances of the order of 1% relative to water in cometary ices. High-spatial resolution sub-mm mapping of HCN distributions supports a nuclear source in some comets (Cordiner et al. 2014). However, it is now known that in many comets a substantial fraction of CN is released from sub-micron dust grains in the inner coma of comets (e.g. Fray et al. 2005).

The C<sub>2</sub> (0-0) band at  $\lambda \approx 5160$  Å may be of comparable flux, but has a lower contrast with the underlying dust grain continuum due to its wider bandwidth. The OH (0-0) band at  $\lambda \approx 3080$  Å generally has higher flux, but unless observed from space the high terrestrial atmospheric absorption severely diminishes the observed flux. Strong CO<sup>+</sup> emission bands have also been observed in a number of comets (the strongest ones are the (3-0) and (2-0) A<sup>2</sup>Π – X<sup>2</sup>Σ; Biver et al. 2018). However this is normally significantly weaker than the observed CN emission, with detection in only less than 20 comets. Given the ubiquitous signature of CN emission in Solar System comets, we have embarked on an archival search for this species in the exocomets around  $\beta$  Pic. Detection of this species would provide a solid link between cometary bodies in our Solar System and the  $\beta$  Pic system.

We detail the observations in Sect. 2, along with the correction for the echelle blaze function and normalization of the spectra to a continuum. Our search for CN relies on a cross correlation search with a theoretical absorption spectrum for the molecule in Sect. 3, and the discussion and interpretation is in Sect. 4 with our conclusions in Sect. 5.

## 2. Data description and analysis

For the study in this paper, we use all the publicly available data from the HARPS and the ESO 3.6m telescope in Chile.

### 2.1. HARPS observations

We collated all publicly available HARPS observations of  $\beta$  Pic obtained between October 2003 and March 2020. Although these observations cover a large baseline, the time between visits varies between 1 day and approximately 2.2 years. Similarly, the number of observations per epoch spans a wide range, between one and 356 exposures, while the exposure times vary between 12 and 900 seconds. This results in an average SNR across order 0 - 10 (order numbers for HARPS will follow the labelling in the fits header – corresponding to 3780 Å and 4075 Å) per exposure between 1 and 430, and a nightly SNR between 63 and 2000. To avoid potential edge effects from order merging, we used the E2DS files downloaded directly from the ESO archive for our analysis.

## 2.2. Initial data processing

The analysis was started from the 2d extracted spectra (E2DS files). Each order was treated separately, with the main focus on orders 3 and 7 as these orders contain the main CN bandhead as well as the Ca II H line, which is used to identify the exocomets. All the spectra were transformed into the stellar rest-frame and onto a common wavelength grid using linear interpolation.

After all the spectra were all transformed into the stellar rest frame, an average nightly spectrum was made from all spectra within a night for each of the nights of observations. This was done after correcting for variations in the effective blaze profile (for example due to instrumental changes, differential slit or fibre losses, atmospheric absorption, etc.), as outlined below. Note that after this correction, the data still have the overall instrumental blaze-profile included.

## 2.3. Effective Blaze function correction

The night of 06 May 2018, which has a high signal-to-noise ratio, was chosen to act as an initial estimate for the blaze function. The individual spectra from this night were averaged together to further boost the signal-to-noise ratio. Subsequently, each individual spectrum was divided by this reference spectrum to highlight the changes in the effective profile. This ratio spectrum was subsequently binned using a bin-width of 81 pixels using a robust algorithm to remove outliers within each bin, clipping points more than 2.7 Median Absolute Deviations (MAD) away from the median of each bin. This binned ratio was then fitted using a second order polynomial. The original spectrum was finally divided by this polynomial to correct for the changes. Note that the change of fibre for HARPS in June 2015 (JD~2,457,180) resulted in an overall change in the blaze profile, and therefore, the spectra before and after this date were separately post-processed as outlined below in Sect. 2.5.

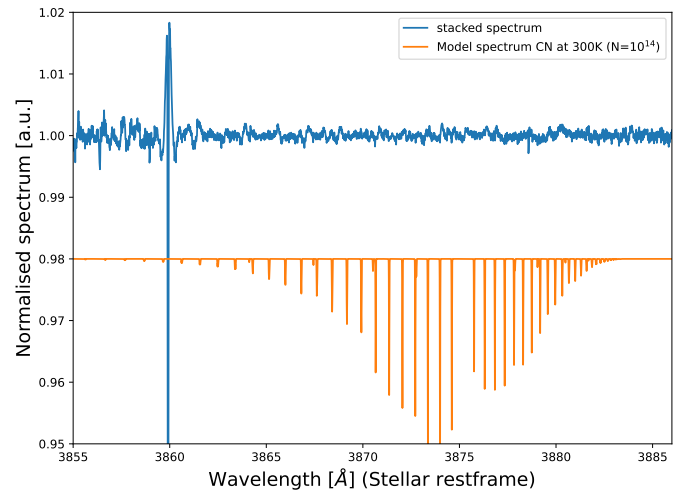
## 2.4. Combining the spectra and flattening them

After the blaze-variation corrections, all the spectra within a night are combined in order to further boost the signal-to-noise and improve the detection limits. Each individual spectrum was weighted using the average signal-to-noise ratio provided in the header for orders 0 and 11 (3780Å - 4075Å).

## 2.5. Removing the stellar spectrum

The processing up to this stage removes most of the variations in the effective blaze function, but the overall stellar spectrum and shape of the blaze profile, are still present and need to be removed. We first perform a second iteration of blaze corrections using a stellar spectrum constructed by averaging the spectra from periods with a low comet activity (see Sect. 3.1), similar to the method before, however, now using a wider binsize of 150 pixels and clipping outliers at 4.5 MAD away from the median. This is done separately pre- and post-fibre change. The correction is done similar to Sect. 2.3, however, an 8th order polynomial was used.

After this post-processing the overall (residual) blaze profile and stellar spectrum needs to be removed. Care must be taken to prevent the removal of (potential) narrow features from the circumstellar disk (as seen in, for example, the Fe and Ca II lines). This is done by combining the nightly spectra without significant cometary activity (see Sect. 3.1) and binning this stacked spectrum in wavelength in bins of 21 pixels. Before binning, the spec-



**Fig. 1.** Plot of the stacked spectrum in the stellar restframe (blue) with a CN model for  $T = 300$  K and  $N = 10^{14} \text{ cm}^{-2}$  overplotted in orange. The large signal around 3860Å is due to the residual from a circumstellar Fe I line.

trum in each bin was median absolute deviation (MAD) clipped to within  $\pm 4.5$  MAD of the median of the flux in the bin. The clipped average of the fluxes within each bin was used. Subsequently, the overall spectrum of the star and (residual) blaze was determined by cubic interpolation across the binned points, and divided out of the spectrum for each night. As before, this was done separately pre- and post-fibre change.

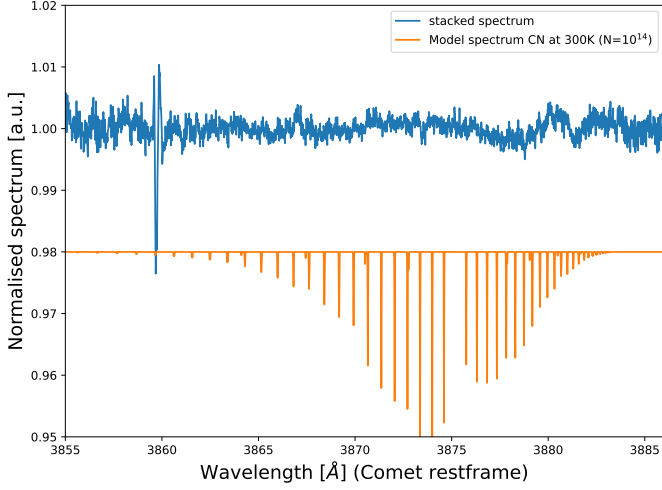
## 3. Searches for CN

After the initial data analysis, we proceed to with our search for lines from CN. Although we are interested in lines from the exocomets, we first co-add the data from all the individual epochs, in the stellar rest-frame, to obtain a deep search for circumstellar CN absorption. These coadded spectra are shown in Fig. 1. No obvious feature from the CN band jumps out, and we need to combine the signal from the lines within the CN band to improve the detection limits. Since the exocomets have a clear radial velocity with respect to  $\beta$  Pic, and since each exocomet will have its own distinct velocity, we selected the 36 epochs with the strongest exocometary features, as seen in the Ca II H line.

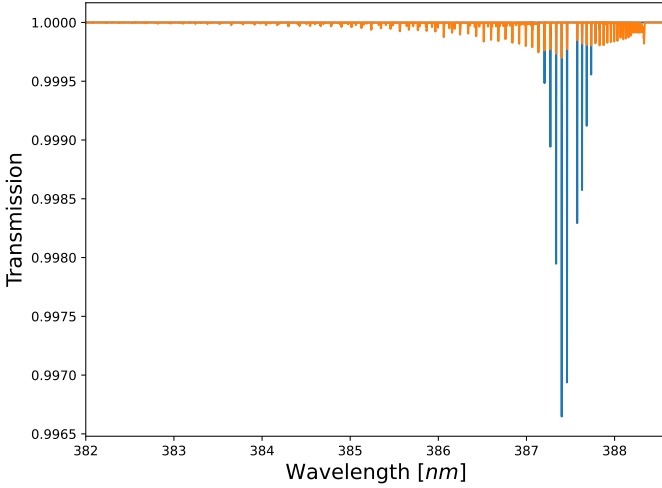
### 3.1. Identifying exocomets from Ca II H line

The comet features in the Ca II H-line ( $\lambda = 3968 \text{ Å}$ ) were identified after removing the overall stellar line-profile and avoiding the absorption from the circumstellar disk by excluding the central  $8 \text{ km s}^{-1}$ . The deepest absorption feature in each observing night was selected, and the location and depth stored. For the comet sample we selected features deeper than 60% while spectra with features below 15% were flagged as ‘low activity’. The distribution of feature locations and depths is shown in Fig. 5. The dates, velocities and depths of the selected exocomets can be found in Tab. A.1

The spectra for the strong comet sample were coadded, and the combined spectrum for the HARPS data is shown in Fig. 2. As with the circumstellar CN, there are no obvious lines from CN visible, and we will need to combine multiple lines to improve the signal-to-noise ratio, and see if there is any CN present.



**Fig. 2.** Similar to Fig. 1 but spectra are stacked in the rest frame of the comet. The large signal around 3860 Å is due to the residual from a circumstellar Fe I line.



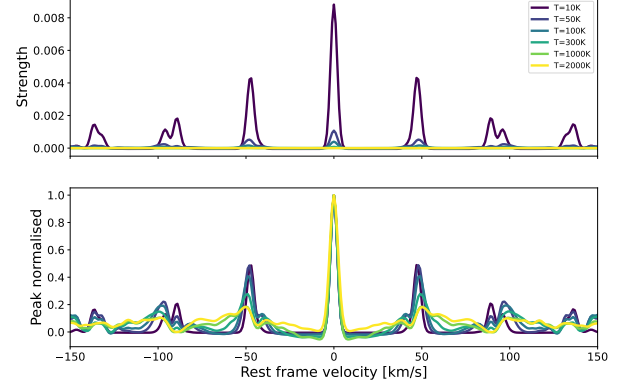
**Fig. 3.** Model absorption spectrum for CN at 30 K (blue line) and 2000 K (orange line). The column densities are  $N = 10^{12} \text{ cm}^{-2}$  and  $N = 10^{13} \text{ cm}^{-2}$ , respectively.

### 3.2. Cross-correlation with CN models

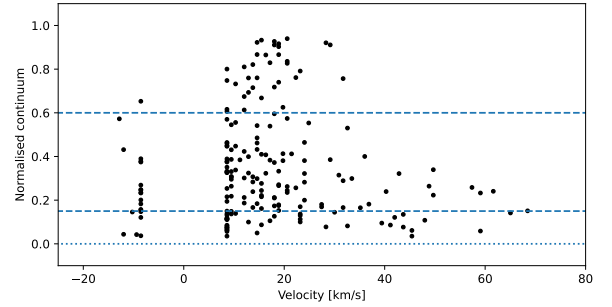
To combine the signal from the many CN lines, we cross-correlate the data for each night with a grid of model of CN. Cross-sections for CN were obtained from ExoMol<sup>1</sup> (Tennyson & Yurchenko 2012; Tennyson et al. 2016; Brooke et al. 2014) with molecular energy levels assumed to be in thermal equilibrium, i.e. Boltzmann distributed. We made models for a wide range of temperatures ( $T = \{10, 20, 30, 50, 100, 200, 300, 500, 1000, 2000, 3000\} \text{ K}$ ) and column densities<sup>2</sup> ( $N = \{10^{11} \dots 10^{15}\} \text{ cm}^{-2}$ ). Each of the models was then convolved to include an effective velocity dispersion using a Gaussian for  $\text{FWHM} = \{2.9, 5, 10, 15, 20\} \text{ km s}^{-1}$ . A value of  $2.9 \text{ km s}^{-1}$  corresponds approximately to the instrumental resolution of HARPS. The theoretical spectra of CN for several different temperatures are shown in Fig. 3.

<sup>1</sup> <https://www.exomol.com>

<sup>2</sup> Note that technically, for optically thin lines, this assumes a filling factor  $f = 1$ . Since  $f$  is unknown, we actually measure  $Nf$



**Fig. 4.** CN autocorrelation function for different temperatures. The aliasing at  $\sim 50 \text{ km s}^{-1}$  is clearly visible



**Fig. 5.** Peaks and radial velocities of exocomets identified in the stellar spectrum. The peaks are normalised such that 1 represents total absorption in the spectrum. Peaks closer than  $10 \text{ km s}^{-1}$  are masked out due to contamination from the circumstellar material. Peak values above 0.6 are counted as exocomets and those below 0.15 as stellar only.

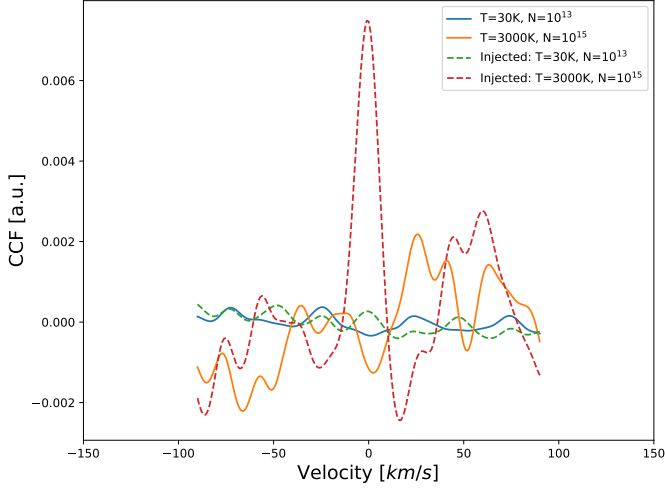
When creating the CCFs, the region around the narrow, circumstellar Fe I lines in this wavelength region (see Figs. 1 and 2) were masked to avoid contamination, especially for the higher temperature CN models. We masked the lines detected in (Kiefer et al. 2019), using a width of  $\pm 0.35 \text{ Å}$  for the Fe I 3859.9 Å line, while for the other lines a width of  $\pm 0.15 \text{ Å}$  was used. Even after cross-correlating, the stacked CCFs (see Fig. 6) do not show any clear signals, and we therefore proceed to estimate our detection limits.

### 3.3. Estimating the CN column density limits

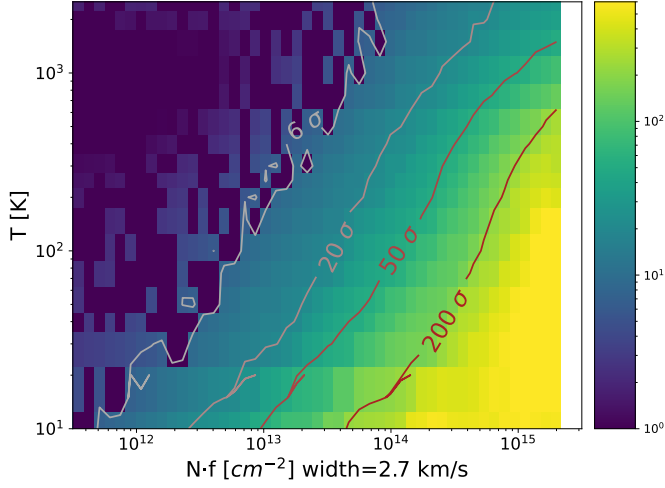
To estimate the limits on the CN detection, we perform injection and recovery tests. We injected models for different temperatures, column densities and velocity broadening into the nightly stacked spectra (Sect. 2.4), and then processed them in the same way as the original data (i.e., correct for residual blaze variations, removal of the stellar envelope and cross-correlation). This was done for all spectra with the injected signal in the stellar rest frame as well as for the sample of spectra with strong cometary activity in the rest frame of the strongest exocometary feature per night.

The individual CCFs show some overall slopes and offsets, and we fit each individual CCF with a combination of a second-order polynomial baseline and a Gaussian with the position fixed to  $0 \text{ km s}^{-1}$  in its rest frame. The fit has five free parameters:





**Fig. 6.** Cross correlation for different number densities of CN averaged over many spectra.



**Fig. 7.** CN measurement limits in the stellar rest frame for a CN line width of 2.7 km s<sup>-1</sup>. Contours show signal-to-noise upper limits.

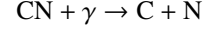
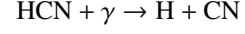
the amplitude of the peak, the width of the peak, and the offset, slope, and intercept of the baseline. To avoid the aliasing peaks of the CN band (see Fig. 4), we only fit the central  $\pm 35$  km s<sup>-1</sup> of the CCF. Fitting is done using the `curve_fit` function from SciPy (Virtanen et al. 2020). We correct each of the individual CCFs for the baseline and then stack the corrected CCFs to obtain our master CCF that we use for determining the upper limit. The noise on each point in the master CCF is determined by taking the standard deviation in time for each pixel and dividing that by the square root of the number of points.

This master CCF is then fitted in the same way as the nightly CCFs, and the uncertainties on the parameters is taken directly from `curve_fit` assuming no covariance between the parameters. The final detection significance for the injected models is determined by the amplitude divided by its uncertainty, and shown in Fig. 7 for the circumstellar gas and in Fig. 8 for the strongest exocomet features.

## 4. Discussion

As discussed above in the introduction, most studies of Solar System comets observe the CN molecules via optical fluorescence. From measurements and simple physical arguments the CN absorption or emission is optically thin everywhere. This means a flux from part of the coma can immediately be converted to a column density using a photon scattering efficiency for the molecular band.

CN in Solar System comets comes from two sources. Sublimation of nucleus near-surface ices releases HCN, which then photo-dissociates into CN. This in turn eventually photo-dissociates into the component atoms:



Since observations of Comet Halley (A’Hearn et al. 1986), it has also been known that dust grains ejected from the nucleus also directly release CN into the coma. However, CN production in comets is usually dominated by nuclear ice sublimation (Fray et al. 2005). The resulting CN coma will have a true space density and apparent column density distribution centrally peaked at the comet nucleus in this case.

For Solar System comets, measured CN column densities are often converted to an overall production rate by the comet of  $Q_{\text{CN}}$  molecules per second via an analytical model known as the “Haser” model (Haser et al. 2020). This model gives the on-sky column density as a function of projected radial distance from the nucleus using only three parameters; the effective destruction scale length  $l_p$  of the parents of CN (HCN+dust), the photodissociation scale length  $l_{\text{CN}}$  of CN, and the gas outflow velocity  $v$ . The Haser model is somewhat simplistic but it represents the data well for Solar System comets observed at typical distances from the Sun (1–3 au).

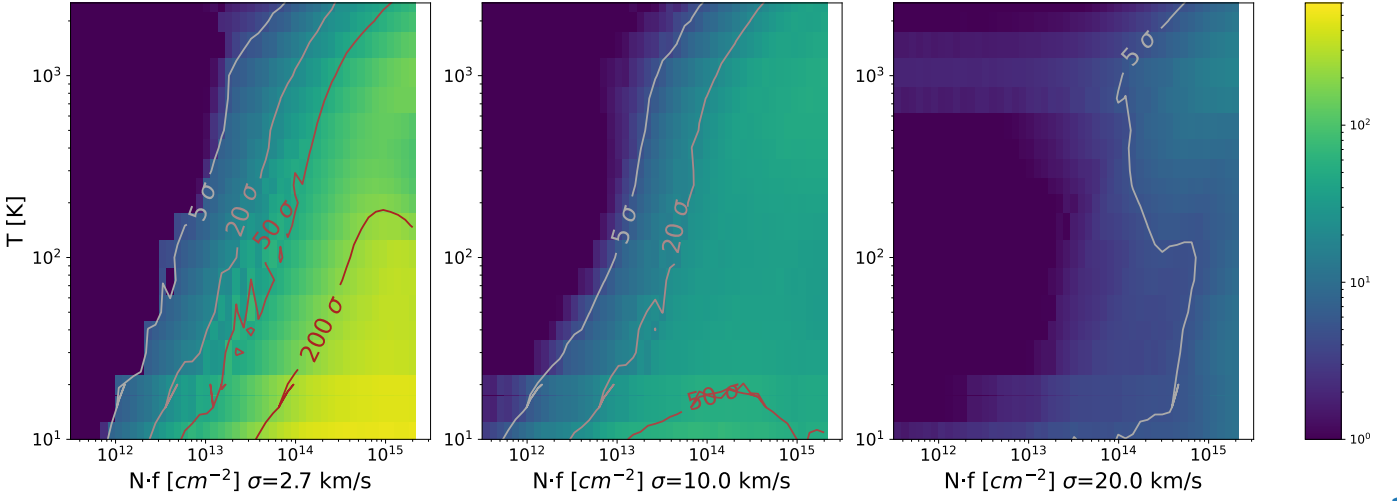
### 4.1. Example comet - 1P/Halley

We take the observed activity of 1P/Halley in 1985/1986 as a reference point for our estimates. Although most Solar System comets are observed to be far less active than Halley, some Long-Period comets show activity in excess of Halley. Given that we will be dominated by the most active comets at  $\beta$  Pic, this seems reasonable as something that can be later scaled to the most active comets such as Hale-Bopp (Schleicher et al. 2024). When Halley was approximately one au from the Sun, it was measured to have  $Q_{\text{CN}} \approx 10^{27}$  molecules per second using Haser models (A’Hearn et al. 1995). The assumed outflow velocity in the models was generally 0.85 – 1 km s<sup>-1</sup> at this distance. For Solar System comets, the normal Haser scale lengths used are (Cochran & Barker 1986; A’Hearn et al. 1995):

$$l_p = 1.7 \times 10^4 r_h^2 \text{ km}$$

$$l_{\text{CN}} = 2 - 3 \times 10^5 r_h^2 \text{ km}$$

Both are assumed to scale with heliocentric distance  $r_h^2$  as photodissociation or thermal desorption is assumed the main physical pathway for both the parent and CN. The lifetime of HCN for photodissociation into H+CN at 1 au has been calculated to be  $(3.2 - 7.7) \times 10^4$  sec, with lifetime of CN dissociating into C+N being  $(1.3 - 3.1) \times 10^5$  sec (Huebner et al. 1992), depending on solar activity levels. These are consistent with the measured scale lengths above if the parents of CN are ejected at approximately 0.5 km s<sup>-1</sup>, and we take into account the excess



**Fig. 8.** CN measurement limits in the exocomet frame for a CN line width of 2.7, 10 and 20 km s<sup>-1</sup>. Contours show signal-to-noise upper limits.

photodissociation energy giving CN an extra kick of  $\approx 0.9$  km s<sup>-1</sup> (Fray et al. 2005).

#### 4.2. 1P/Halley at $\beta$ Pic

Photodissociation is driven by the UV photon flux at the comet, which in turn is dominated by the Ly- $\alpha$  line flux (Combi & Delsemme 1980). The Ly- $\alpha$  emission from  $\beta$  Pic has been measured by Wilson et al. (2017) via HST+COS. Using their ISM and CSM corrected Voigt line profile parameters, we integrated over  $\pm 600$  km s<sup>-1</sup> and corrected for the distance of  $\beta$  Pic to derive an intrinsic Ly- $\alpha$  luminosity of  $1.4 \times 10^{30}$  ergs sec<sup>-1</sup>.

The solar Ly- $\alpha$  flux varies by  $\sim 50\%$  over the Solar cycle, with a mean value at Earth of  $4.7 \times 10^{11}$  phot s<sup>-1</sup> cm<sup>-2</sup> (Woods et al. 2000). The resulting mean solar Ly- $\alpha$  luminosity is  $2.1 \times 10^{28}$  ergs sec<sup>-1</sup>, a factor of  $\sim 67$  lower than  $\beta$  Pic. Hence the photodissociation lifetime of CN at 1 au from  $\beta$  Pic of an ‘exoHalley’ is approximately 67 times shorter, for example around 3000 sec.

In steady-state sublimation, the sublimation rate is proportional to the photon flux at the comet. For  $\beta$  Pic  $L = 8.7L_{\odot}$ , so assuming all CN production scales with stellar luminosity, an exoHalley has a larger CN production rate of  $Q_{\text{CN}} \sim 9 \times 10^{27}$  molecules per second. The total number of CN molecules in the coma at 1 au at any given time would therefore be:

$$N_{\text{CN}} = Q_{\text{CN}} \tau_{\text{CN}}$$

$$N_{\text{CN}} \sim 9 \times 10^{27} \cdot 3000 \sim 3 \times 10^{31} \text{ molecules}$$

The effective cross section of the coma will depend on how much is projected against the disk of  $\beta$  Pic. The radius of the CN coma is then given by

$$r \sim v_{\text{CN}} \tau_{\text{CN}}$$

The outflow velocity of the parent HCN will depend on the nuclear surface temperature, which will increase as a comet moves towards its parent star. Biver et al. (2002) found a nuclear gas outflow velocity  $v \approx 1.0 r_h^{-0.4}$  km s<sup>-1</sup> from sub-mm measurements of Comet Hale-Bopp, close to the canonical relationship of  $r_h^{-0.5}$ . Harris et al. (2002) notes lower activity similar to Halley results in smaller outflow velocities. The excess photodissociation velocity of CN will not change. Therefore gas outflow

speeds are expected to be  $\sim 1 - 2$  km s<sup>-1</sup> at 1 AU, and the CN coma radius will be:

$$r \sim (1 - 2) \times 3000 \leq 6 \times 10^3 \text{ km.}$$

This is much smaller than a stellar radius, so we can assume for an exoHalley that all of the coma is projected into the stellar disk as seen from Earth.

$\beta$  Pic has  $R = 1.8R_{\odot} \equiv 1.3 \times 10^{11}$  cm. Given that the coma is optically thin, and we do not resolve the star, the effective average column density towards  $\beta$  Pic is:

$$n \sim \frac{N_{\text{CN}}}{\pi R_*^2}$$

$$n \sim \frac{3 \times 10^{31}}{\pi \cdot 1.7 \times 10^{22}} \sim 6 \times 10^8 \text{ cm}^{-2},$$

therefore an exoHalley at 1 au from  $\beta$  Pic would result in a CN column density of  $10^8 - 10^9$  cm<sup>-2</sup>. This is  $\geq 100$  times smaller than our measured upper limits above.

#### 4.3. Raising the CN production rates

There are several ways exocomets at  $\beta$  Pic may exhibit larger CN production rates. A larger comet gives more surface area for sublimation. Comet Halley’s nucleus has an effective radius of 5.5 km (Keller et al. 1987), whereas Comet Hale-Bopp was one of the most intrinsically active comets on record with an extremely large nucleus of radius around 14 – 21 km (Weaver et al. 1997). However most of the apparent brightness was due to dust emission, and its only had  $Q_{\text{CN}} \approx 3 \times 10^{27}$  molecules per second (Schleicher et al. 2024). In this case, we might anticipate the most active comets in the Solar System to have  $Q_{\text{CN}} \sim 10^{28}$  molecules per second, implying a commensurate increase of a factor 10 in column density  $n$  over an exoHalley at  $\beta$  Pic.

Most comets only have a fraction of their surface area undergoing mass-loss, due to the effect of a thermally insulating dust crust. Halley itself only has  $\sim 20\%$  of the sunlit surface active (Keller et al. 1987). Yet some comets are known as hyperactive, as the measured gas productions rates would require  $\geq 100\%$  of the nucleus surface to be active (Sunshine & Feaga 2021). This is believed to be due to the release of ice-rich dust grains from the nucleus, which sublimate and act as a secondary source of gas.

Assuming this is possible for exocomets at  $\beta$  Pic would increase  $N_{\text{CN}}$  and  $n$  by a factor of around five.

The closer a comet to its parent star, the greater the heating and corresponding production rate. When insolation completely controls the nuclear surface energy balance we expect  $Q_{\text{CN}} \propto r_h^{-2}$ . Studies of Halley and Hale-Bopp showed power laws close to this (e.g. Biver et al. 2002). But the photodissociation lifetime  $\tau_{\text{CN}} \propto r_h^2$ , so theoretically:

$$N_{\text{CN}} = \frac{Q_{\text{CN}}(1 \text{ au})}{r_h^2} \cdot \tau_{\text{CN}}(1 \text{ au}) r_h^2 = \text{constant},$$

hence the total amount of CN in the coma should be approximately independent of stellar distance, and the total cross-section would not change.

Finally, comets have impulsive outbursts, where the amount of material ejected into the coma increases significantly over a few hours or less. The general causes are unclear, but there is evidence for several physical mechanisms that cause them (Kelley et al. 2021; Müller et al. 2024). “Normal” outbursts generally result in a brightness increases of 1–5 magnitudes, with smaller outbursts being more common. As the apparent brightness is directly proportional to the amount of material (gas+dust) in the coma, this implies short term increases in  $N_{\text{CN}}$  and  $n$  of factors ranging from 2 to 100. Large outbursts are often due to partial or total disruption of a comet and can occur when a comet is closer to the Sun. Hence there could be a correlation with faster moving  $\beta$  Pic exocomets nearer their periastra producing greater measurable column densities of CN.

In summary, mechanisms exist by which the CN column density of an exocomet at  $\beta$  Pic could exceed  $10^9 \text{ cm}^{-2}$ . A larger or more hyperactive nucleus could increase  $n$  by a factor of ten. Even larger increases are possible with an outburst or comet disruption, but the upper boundary is likely to be near  $n \sim 10^{11} \text{ cm}^{-2}$ .

## 5. Conclusions

We have combined over 1000 spectra of  $\beta$  Pic over a 20 year period, identified two groups, one with deep exocomet signals, and another to act as a stellar reference. We produce a normalised spectrum towards  $\beta$  Pic with the stellar spectral features removed, and then aligned the spectra to the exocomet rest frames and perform a search for CN using cross correlation.

We do not detect CN both in the stellar rest frame and in the exocomet rest frame. Injection and recovery tests show that for the circumstellar CN we are sensitive to  $10^{12.5} \text{ cm}^{-2}$  below 80 K and below  $10^{14} \text{ cm}^{-2}$  up to 1800 K. In the exocomet rest frames, sensitivity is somewhat similar. Below 100 K the sensitivity is  $10^{12} \text{ cm}^{-2}$  and above from 100 - 1800 K, is  $10^{13} \text{ cm}^{-2}$ . This is a factor of 10 to 100 too high to detect a single Hale-Bopp equivalent comet with a CN coma. The assumptions are for steady state comets, so individual outbursts could build this number up, and multiple transiting exocomets can increase the absorption cross section.

$\beta$  Pic is the brightest star with known exocomets seen both in transit and spectroscopy, and represents the best candidate for exocomet characterisation. A more dedicated observing campaign over several years using dedicated spectrographs can reach the equivalent Solar System limits for CN. PLATO will carry out a dedicated long stare campaign that will observe the star with sub-millimagnitude photometry over two to three years. This presents an excellent opportunity for coordinated spectroscopic campaigns where the exocomet population of  $\beta$  Pic can be

characterised, with CN and ideally other molecules known from Solar System comets to be identified in another stellar system.

## Data availability

An online repository with materials used in this work is available at <https://github.com/mkenworthy/BetaPicCN> using the `showyourwork!` package (Luger et al. 2021).

**Acknowledgements.** The authors thank the Lorentz Centre at Leiden University for organising the workshop “Exocomets: Understanding the Composition of Planetary Building Blocks” which led to the study in this paper. Part of this work was supported by the German *Deutsche Forschungsgemeinschaft*, DFG project number Ts 17/2–1. AF acknowledges support from STFC award ST/T00021X/1. EdM acknowledges support from STFC award ST/X00094X/1. We made use of the Python programming language (Rossum 1995) and the open-source Python packages `numpy` (van der Walt et al. 2011), `matplotlib` (Hunter 2007), `astropy` (Astropy Collaboration et al. 2013).

## References

- A’Hearn, M. F., Hoban, S., Birch, P. V., et al. 1986, *Nature*, 324, 649
- A’Hearn, M. F., Millis, R. C., Schleicher, D. O., Osip, D. J., & Birch, P. V. 1995, *Icarus*, 118, 223
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Bendahan-West, R., Kennedy, G. M., Brown, D. J. A., & Ström, P. A. 2024, *MNRAS*[arXiv:2412.13253]
- Beust, H., Lagrange, A. M., Plazy, F., & Mouillet, D. 1996, *A&A*, 310, 181
- Beust, H., Vidal-Madjar, A., Lagrange-Henri, A. M., & Ferlet, R. 1991, *A&A*, 241, 488
- Biver, N., Bockelée-Morvan, D., Colom, P., et al. 2002, *Earth Moon and Planets*, 90, 5
- Biver, N., Bockelée-Morvan, D., Paubert, G., et al. 2018, *A&A*, 619, A127
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977
- Brooke, J. S. A., Ram, R. S., Western, C. M., et al. 2014, *ApJS*, 210, 23
- Bus, S. J., A’Hearn, M. F., Schleicher, D. G., & Bowell, E. 1991, *Science*, 251, 774
- Cochran, A. L. & Barker, E. S. 1986, in *ESA Special Publication*, Vol. 250, *ESLAB Symposium on the Exploration of Halley’s Comet*, ed. B. Batrick, E. J. Rolfe, & R. Reinhard, 439
- Combi, M. R. & Delsemme, A. H. 1980, *ApJ*, 237, 641
- Cordiner, M. A., Remijan, A. J., Boissier, J., et al. 2014, *ApJ*, 792, L2
- Cremonese, G., Boehnhardt, H., Crovisier, J., et al. 1997, *ApJ*, 490, L199
- Cremonese, G., Huebner, W. F., Rauer, H., & Boice, D. C. 2002, *Advances in Space Research*, 29, 1187
- Dufay, J., Swings, P., & Fehrenbach, C. 1965, *ApJ*, 142, 1698
- Ferlet, R., Hobbs, L. M., & Madjar, A. V. 1987, *A&A*, 185, 267
- Fitzsimmons, A. & Cartwright, I. M. 1996, *MNRAS*, 278, L37
- Fray, N., Bénilan, Y., Cottin, H., Gazeau, M. C., & Crovisier, J. 2005, *Planet. Space Sci.*, 53, 1243
- Fulle, M., Leblanc, F., Harrison, R. A., et al. 2007, *ApJ*, 661, L93
- Fulle, M., Molaro, P., Buzzi, L., & Valisa, P. 2013, *ApJ*, 771, L21
- Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, *AJ*, 132, 161
- Harris, W. M., Scherb, F., Mierkiewicz, E., Oliverson, R., & Morgenthaler, J. 2002, *ApJ*, 578, 996
- Haser, L., Oset, S., & Bodewits, D. 2020, *Planetary Sciences Journal*, 1, 83
- Heller, R. 2024, *A&A*, 689, A97
- Hobbs, L. M., Vidal-Madjar, A., Ferlet, R., Albert, C. E., & Gry, C. 1985, *ApJ*, 293, L29
- Huebner, W. F., Keady, J. J., & Lyon, S. P. 1992, *Ap&SS*, 195, 1
- Hui, M.-T. 2023, *AJ*, 165, 94
- Hunter, J. D. 2007, *Computing In Science & Engineering*, 9, 90
- Hyland, M. G., Fitzsimmons, A., & Snodgrass, C. 2019, *MNRAS*, 484, 1347
- Keller, H. U., Delamere, W. A., Reitsema, H. J., Huebner, W. F., & Schmidt, H. U. 1987, *A&A*, 187, 807
- Kelley, M. S. P., Farnham, T. L., Li, J.-Y., et al. 2021, *Planetary Sciences Journal*, 2, 131
- Kennedy, G. M. 2018, *MNRAS*, 479, 1997
- Kennedy, G. M., Hope, G., Hodgkin, S. T., & Wyatt, M. C. 2019, *MNRAS*, 482, 5587
- Kiefer, F., Lecavelier des Etangs, A., Boissier, J., et al. 2014, *Nature*, 514, 462
- Kiefer, F., Lecavelier des Etangs, A., Vidal-Madjar, A., et al. 2017, *A&A*, 608, A132
- Kiefer, F., Vidal-Madjar, A., Lecavelier des Etangs, A., et al. 2019, *A&A*, 621, A58

- Koen, C., Balona, L. A., Khadaroo, K., et al. 2003, MNRAS, 344, 1250
- Lagrange-Henri, A. M., Gosset, E., Beust, H., Ferlet, R., & Vidal-Madjar, A. 1992, A&A, 264, 637
- Lecavelier des Etangs, A., Cros, L., Hébrard, G., et al. 2022, Scientific Reports, 12, 5855
- Lecavelier Des Etangs, A., Vidal-Madjar, A., & Ferlet, R. 1999, A&A, 343, 916
- Lee, R. A., Gaidos, E., van Saders, J., Feiden, G. A., & Gagné, J. 2024, MNRAS, 528, 4760
- Lindgren, L., Klioner, S. A., Hernández, J., et al. 2021, A&A, 649, A2
- Luger, R., Bedell, M., Foreman-Mackey, D., et al. 2021, arXiv e-prints, arXiv:2110.06271
- Müller, D. R., Altwegg, K., Berthelier, J.-J., et al. 2024, Monthly Notices of the Royal Astronomical Society, 529, 2763
- O’Brien, D. P., Izidoro, A., Jacobson, S. A., Raymond, S. N., & Rubie, D. C. 2018, Space Sci. Rev., 214, 47
- Pavlenko, Y., Kulyk, I., Shubina, O., et al. 2022, A&A, 660, A49
- Pollard, K. 2024, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 56, AAS/Division for Extreme Solar Systems Abstracts, 621.12
- Preston, G. W. 1967, ApJ, 147, 718
- Rappaport, S., Vanderburg, A., Jacobs, T., et al. 2018, MNRAS, 474, 1453
- Rossum, G. 1995, Python Reference Manual, Tech. rep., Centrum voor Wiskunde en Informatica (CWI), Amsterdam, The Netherlands, The Netherlands
- Schleicher, D. G. 2010, AJ, 140, 973
- Schleicher, D. G., Birch, P. V., Farnham, T. L., & Bair, A. N. 2024, Planetary Sciences Journal, 5, 281
- Schmidt, C. 2016, Icarus, 265, 35
- Slaughter, C. D. 1969, AJ, 74, 929
- Sunshine, J. M. & Feaga, L. M. 2021, Planetary Sciences Journal, 2, 92
- Tennyson, J. & Yurchenko, S. N. 2012, MNRAS, 425, 21
- Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, Journal of Molecular Spectroscopy, 327, 73, new Visions of Spectroscopic Databases, Volume II
- Thackeray, A. D., Feast, M. W., & Warner, B. 1966, ApJ, 143, 276
- Tobin, W., Barnes, S. I., Persson, S., & Pollard, K. R. 2019, MNRAS, 489, 574
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22
- Vidal-Madjar, A., Hobbs, L. M., Ferlet, R., Gry, C., & Albert, C. E. 1986, A&A, 167, 325
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261
- Weaver, H. A., Feldman, P. D., A’Hearn, M. F., & Arpigny, C. 1997, Science, 275, 1900
- Wilson, P. A., Lecavelier des Etangs, A., Vidal-Madjar, A., et al. 2017, A&A, 599, A75
- Woods, T. N., Tobiska, W. K., Rottman, G. J., & Worden, J. R. 2000, J. Geophys. Res., 105, 27195
- Zieba, S., Zwintz, K., Kenworthy, M. A., & Kennedy, G. M. 2019, A&A, 625, L13



## Appendix A: Selected exocomets

The dates for the nights selected to have the strongest exocomet features, as well as the depths and velocities of these features can be found in Table A.1.

**Table A.1.** Nights with the strongest exocomet features.

BJD-2,400,000 <sup>a</sup>	$v_{comet}$ [km/s]	Depth
53031.2	19.7	0.63
53032.2	18.9	0.74
53033.1	22.3	0.76
53036.1	15.5	0.67
53045.1	18.0	0.72
53046.1	18.0	0.91
53047.1	20.6	0.84
53048.0	20.6	0.83
53049.1	20.6	0.94
53050.1	18.9	0.90
53051.1	18.9	0.87
53052.1	18.9	0.92
53053.1	18.0	0.93
53054.1	17.2	0.83
53056.1	16.3	0.86
53057.1	15.5	0.93
53058.1	14.6	0.92
53059.1	14.6	0.87
53060.1	13.7	0.82
53061.1	13.7	0.72
53062.1	12.9	0.69
53063.1	12.0	0.67
53064.1	12.0	0.61
53581.4	-8.5	0.65
54223.0	8.6	0.62
54542.1	29.2	0.91
54544.1	28.3	0.92
54545.0	23.2	0.79
54546.0	8.6	0.80
54799.2	8.6	0.75
54878.1	14.6	0.76
54913.0	12.0	0.81
56524.4	31.7	0.76
56534.3	12.9	0.76
56536.3	8.6	0.61
57849.0	10.3	0.73

**Notes.** <sup>(a)</sup> Dates given are for the average of all the observations that were combined during each of the nights.