13/07/2018

Mr Jason Neal

jason.neal@astro.up.pt

Our Ref. : AA/2018/33452

Dear Mr Neal,

Your paper "Towards the near-infrared detection of brown dwarf companions" was submitted to a referee who recommends rejection on the grounds that its scientific content is not sufficient to warrant publication in A&A.

As you can see the main issue from the referee is that this kind of study has been successfully carried out by other groups so your "exploratory" and unsuccessful study, limited by problems with fitting stellar models, is not a useful addition to the literature as it is exploring an area and providing null results, where others have already succeeded. The referee is thus also not convinced that any revisions could make this work acceptable.

Given the nature of the referee’s report, I regret to inform you that I shall be unable to give any further consideration to this paper.

I am sorry to disappoint you on this occasion.

Yours sincerely,

Eline Tolstoy

A&A Editor

------------------------

Referee report

General comments:

1. The stated goal of this paper was to detect the spectra of substellar companions around FGK stars in order to measure their masses. The authors initially wanted to explore the direct subtraction method, but the observations were not taken with a long enough time baseline in between. While it is unfortunate observing specifications were not met, it appears that this technique would not have worked for some of the sample anyway. Given the long orbital periods of some of the targets in their sample, a number of these targets would not have been conducive to this technique even if their observations had been taken as specified. The setup of this project was flawed and preventable at the proposal stage.

2. With the second proposed method (chi2 method), direct constraints on companion mass are no longer possible, so the motivation presented for this study is no longer applicable. The focus/motivation of the paper needs to change. However, it is unclear what the goal of the chi2 method is. Is it to detect the presence of a BD companion that has not been discovered yet? If so there are more straightforward ways to do this (i.e. RVs, astrometry, imaging). Furthermore, it is unclear what the science goal of the chi2 method is. Even if you could detect the companion, what is the science that you would get out? This is not touched on.

3. The authors were not able to detect any of the companions. The results of this study are limited by model fits to the stellar spectra. It is unclear why decent fits are not possible, since given the star/companion contrast you are basically fitting a model atmosphere to a FGK stellar spectrum, which has been done. The issues this study runs into are not even about recovering the companion, but are instead limited by this fit to the stellar spectrum. Because of this limitation, it is not possible to accurately determine how effective the chi2 method is at detecting a companion. Other studies have been able to successfully detect a faint companion (i.e. Snellen et al 2010, Piskorz et al 2016, Birkby et al 2017), in addition to producing science from this detection (i.e. molecule detections, mass measurements, etc.).

4. Concluding by saying this was a good exploratory study is not accurate – mistakes regarding small delta RVs were avoidable, and other studies have explored and successfully achieved detecting faint companions and producing science from those detections. It is unclear what this study adds to the literature.

Comments in order of appearance in the paper:

**1. How were the seven targets chosen? Are all of these from Sahlmann et al 2011?**

Yes this is stated in the motivation section. I have slightly reworded. Based on what targets were visible in that observing period.

2**. Table 2 – helpful additional columns include age of the system and distance to the system.**

Ages were already provided! I have added distances calculated from GAIA parallax. The value if HD4747 is consistent with the valid published in Crepp et al. 2016.

**3. Table 3 – Expand the caption of Table 3 into a separate paragraph going into detail regarding obtaining the flux ratio. Also list the absolute K-band mag for the star and the companion, as well as what parameters are assumed to get the companion magnitude. Table 3 is also not mentioned before Table 4. Describe what is meant by phase coverage. Why is there no delta RV for HD 4747?**

I have moved Table 3 to appendix, rearrange the last section of appendix to explain RV and flux ratio calculations and to point to the Table 3 values.

HD4747 only has 1 observation. Still need to add the Magnitudes of the host and companions.

**4. Justify filter choice (Ks, Hx5e-2, J) for each object.**

Is there a justification? I feel like some were a mistake (during observation)...

**5. Be more specific when discussing the custom reduction pipeline (Figueira et al 2010), particularly regarding parameters altered between K and H band reductions (just specifying that “parameters” were changed is vague).**

I removed the sentence that some parameters were altered. The only substantial change in the ends was the spline function order used for the tracing fit.

**6. Table 4 – Provide an equation for the RV calculation in the text. Also include columns for companion-star separation (arcsec and AU), as well as distance to the system, integration time, number of exposures, and S/N of spectra.**

Have update the table and moved most of the caption to the text.

I have provided an RV equation in appendix.

Have add a S/N per observation from continuum of detector 2.

Should I include distance, Sahlmann has asini and arel in AU. He only has 2 which have asini or arel in (mas).?

**7. How is the degree for the polynomial fit to the continuum selected for each spectrum and detector? Why not use the same degree for all?**

Because that shape was different for each chip. I reworded to just state we continuum normalize by a fitted polynomial.

**8. The authors indicate challenges with regards to bad pixels and cosmic rays. Try y-shifting then median combining the AB pairs before fitting the trace – this is commonly done to avoid problems like this.**

NO

**9. How many AB pairs typically exhibited these artifacts due to bad pixels/cosmic rays? Explicitly indicate for each dataset.**

14\% Have added numbers per each obseration

**10. Fig. 1 – Why is the artifact so extended in the pink optimal extraction spectrum? It is unclear to me why the optimal extraction would produce such an extended feature. Please show the 2D spectrum after rectification for clarification. How are the bad pixels removed?**

NO. have added that we are not sure why it is extented.

**11. Describe what causes instrument glow. Also previous paragraphs indicated that the extended features were due to bad pixels/cosmic rays. Please be more precise when discussing causes of artifacts in the spectra.**

Removed glow statement

**12. The authors indicate that telluric lines have a Voigt profile, but at their resolution a Gaussian profile is a good assumption. Please demonstrate that this is a good assumption. Other common methods for fitting a telluric line profile include fitting a central Gaussian with a series of satellite Gaussians on either side – can you justify using one model over the other?**

Domninated by instrument profile in the NIR. Seifart 2010.

**13. What is meant by “manually” in the description “identification of telluric and stellar lines performed manually for each spectra”?**

Changed to “by hand”

**14. Why don’t you use a stellar model as well for the wavelength calibration? For a number of the targets the RV of the star should be well constrained, and so it should be possible to use the stellar lines as well as the telluric ones to get a more accurate wavelength calibration.**

I mention this speciifcal piskorz who need to use the stellar model at 2 micron where only telluric is fine elsewhere. Need to cite more references that use both models.

1**5. For obtaining a best-fit telluric model, it seems like a better model could be found by directly fitting abundances, scaling etc to the observed spectra. Or by using PCA (see for example Piskorz et al 2016). Given the issues with observed spectra noted later in the paper, perhaps one of these other methods would yield a better fit.**

**16. How much overlap is there between telluric lines and stellar lines? Could the telluric removal impact the stellar lines (perhaps contributing to the poor model fits)?**

Some in detector 2. Have not added anything to the paper though.

**17. For each of the 7 targets, what was the expected delta RV given the intended time baseline specified in the proposal? Given the long orbital periods for some targets, it seems unlikely that sufficient RV offset would be achieved in a given semester.**

In the original proposal it doesn’t specific expected delta RV….

**18. Table 5 – instead of showing the range of the model grid, show ranges used in the chi2 fits for each system.**

??? not sure what he means here. These are visible in the plots...

1**9. When mentioning the use of MAD, describe what this method does (instead of just citing the source). Also justify the choice of the order of approximation and jump parameter.**

Added that it preforms numerical derivatives. Justified saying we followed the directions of use and indicate what the parameters mean.

**20. Why is the RV of the star left as a free parameter in the chi2 grid fits? For some of the systems where the companion is on a fully resolved orbit, isn’t this precisely known?**

To see if the known value was returned.

**21. Make it clear earlier on that the goal is no longer to measure the mass of the companion – this was never explicitly stated after moving on to the chi2 method.**

Added “The temperature of synthetic spectra fitted to the companion will provide some indication of the companions spectral type, but will not produce a direct mass constraint that was the original aim of this work.”

2**2. When introducing the high dimensionality of the binary model, explicitly list all parameters that are included in that model. In addition, instead of fixing most of the parameters, why not do a 2-step grid search, one coarse, another fine?**

Added the list of parameters.

**23. The authors mention later that an iterative process could be implemented to refine the RV grids – recommend doing this, since the plots (i.e. Fig 3) of the grid space look too coarse.**

The companion is much finer. They may be a little coarse, but I don’t think this will change the results significantly

**24. After applying the chi2 method to a simulated dataset, the authors say that the reduced chi2 for the binary model is closer to 1, indicating that the binary model is a better fit. This is not necessarily the case. A proper model comparison needs to be done, such as using BIC to compare models – as the binary model has more free parameters, it is not surprising that it is a better fit, but this does not mean that model is preferred.**

Added BIC analysis to show they are significant. Just need to find correct m and n for observation run.

**25. For HD 211847, what is the separation of the companion and the star (arcsec)? What is the % flux from the star at the location of the companion? Was the slit oriented along the PA of the system? Do you take any of these factors into account when estimating contrast?**

Separation is 222mas or 11.3 Au from Moutou 2017... Not sure on the slit orientation we had???

**26. Masking out lines in the model spectrum in order to get a better fit does not seem to address the underlying problem. It is unclear to me why a reasonable fit to a stellar spectrum is not possible here, as this is not new. It is also concerning that different detectors yielded different results. This indicates that there might be a problem with the reduction of the spectra.**

I was also got similar limits with synthetic simulations. (no data involved)

**27. Fig. 7 – in the caption, explicitly state that these ± 100 K error bars are from the grid size.**

Done

**28. I think the characterization of a derived upper limit of ~600 M\_J for this method is misleading. At present the analysis seems to be limited by a poor match between stellar models and the stellar spectra, and so it does not seem possible to extrapolate anything about the mass limits of this technique regarding companion detection given a fundamental problem with the analysis occurred upstream.**

I have changed the claim in the abstract, that the limit was from the injection-recovery.

**29. I disagree with the characterization of this study as “exploratory” – detecting faint companions using high-res NIR spectra is not new, and has been thought about by a number of groups (i.e. Birkby et al 2017, Snellen et al 2010, Piskorz et al 2016).**

**30. The authors state that their non-detection of binary companions with low flux ratios is consistent with other works. But this study is limited by model fits to the stellar spectra, not detection of the faint companions. Also, if other studies had trouble detecting faint companions, what was different about this study? What made the authors think that they might be able to detect the faint companions? These points are not made clear.**

**31. The authors note that for systems HD 4747, HD 211847, and HD 202206, the lines of both components are blended. It is unclear how this happened, as this should have been straightforward to plan around for scheduling of observations.**

I have put that it was due to poor planing

**32. At the start of section 7.1.5, it is unclear if the statement regarding the wavelength range being not the best choice was referring to the original plan to do direct subtraction, or for the chi2 method. Please clarify.**

Added some more clarity

**33. It is unclear why the target list was not more thoroughly vetted given that a number of the systems would not have been amenable for the direct subtraction technique that was the original goal of this study.**

From my understanding this was based on what targets were observable in that period.

**34. Conclusions – the authors state that the objective of the observations was to apply the differential technique. But many of their targets would not have worked with this technique due to their long periods, so it is unclear how this was the original goal. While it is unfortunate observing specifications were not met, it appears that this technique would not have worked for a number of their targets anyway. Furthermore, the statement that the spectral recovery method is equivalent is incorrect. The original goal of the study (as motivated in the paper) was to get mass measurements – this is not possible with the chi2 method. Finally, please cite studies that indicate issues with model mismatches. I’ll note that the limitation with this study seems to be getting a good fit to the stellar spectrum – other studies have been able to do this (in particular when trying to detect faint companions), so it does not seem like models are entirely to blame here. There are also better methods for detecting faint companions that have been demonstrated previously (such as the cross-correlation method).**

Added 3 papers that briefly mention line mismatches in with phoneix-aces or bt-settl compared to observations.

Removed that both methods are “fully equivalent” from conclusion aswell as removed the word “good”