**Response to reviewer report: FIESTA II. Disentangling stellar and instrumental variability from exoplanetary Doppler shifts in Fourier domain**

We thank the all the science/data/statistics reviewers for their helpful and insightful comments. All reviewer comments are repeated below in *shadow\_and\_italics*, followed by the response (black) and the changes to the manuscript (purple texts or figures wrapped in purple frame).

1. *Introduction: the CCF is introduced by saying that it is a cross-correlation “of the stellar spectrum with a template spectrum or synthetic mask,” but the type of CCF used throughout this work appears to be specifically the EPRV standard cross-correlation with a weighted binary mask, which produces a CCF with the shape of an average spectral line. It’s not clear how the results & interpretation of the FIESTA method would change with a different kind of CCF. If the authors have thoughts on this, it would be an interesting discussion point! Regardless of that, the introduction should be more specific about the definition of the CCF that’s adopted in the paper, with a reference to Pepe et al. (2002) or a similar paper that goes into more detail.*

We agree that a different choice of CCF mask would inevitably change the CCF shape and thus result in different FIESTA outputs, but this is analogous to the scenario of choosing a different CCF mask to generate a different version of RVs. While it would be interesting to study the effects of different types of masks used to construct the CCF on the FIESTA outputs, whether this would benefit the stellar activity modelling and the planet search remains to be explored and can be in a separate paper. This paper rather focuses on the FIESTA method and uses the simulated SOAP data as well as the HARPS-N solar data to demonstrate the method, so we take the CCFs as they are from SOAP and HARP-N pipeline. However, generally speaking, we do need to be sure that the same mask is used within one FIESTA analysis (e.g., on a star over the period to be studied), as we do not want the CCF shape changes from different CCF mask being used in a set of data.

For clarity, we add the references to the sentence in question.

“In the context of radial velocity detection of exoplanets, the signal is usually the cross-correlation function (CCF) of the stellar spectrum with a template spectrum or synthetic mask (Baranne et al. 1996; Pepe et al. 2002).”

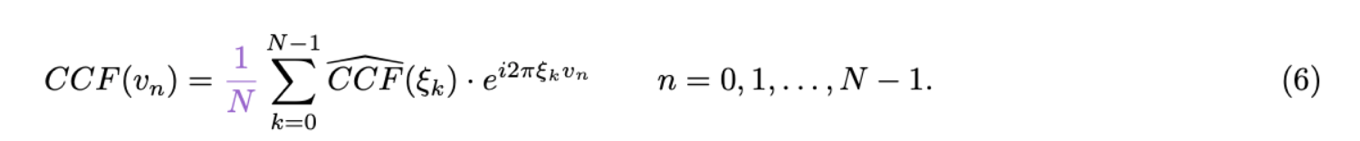
We also add the new Section 7.3.3 for such a discussion.

“In this paper, we demonstrate the *Φ*ESTA methodology in the simulated SOAP data as well as the HARPS-N solar data. The CCFs are taken as they are from the SOAP and HARP-N pipeline.

We are aware that the way the CCF is constructed, whether a different choice of CCF mask or different orders in the échelle spectrograph being used, would inevitably change the CCF shape and thus result in different *Φ*ESTA outputs. While it would be interesting to study the effects of different forms of CCFs generated from the same stellar spectra on the *Φ*ESTA outputs, whether this would benefit the stellar activity modelling and the planet search remains to be explored. However, we do need to be sure that the CCFs used within one *Φ*ESTA analysis (e.g., on a star over the period to be studied) are constructed in a consistent way, as we do not want the CCF shape changes from the data reduction process. For the same reason, the CCF continuum normalisation needs to be robust.”

1. *Section 2.1: Equation 6 may be missing a factor of 1/N.*

Indeed. 1/*N* was missing and is now added.



Meanwhile, we notice that the discrete Fourier transform in Equation 2 has an extra term 1/*N*, which is now removed.

1. *Figure 1 is not referenced in the text. It’s a very useful figure for building understanding of the method and should definitely be discussed!*

Figure 1 was first referenced in Section 3.1, but we realise that it was inadequately explained in the text, especially the left half of the original figure, which is the current (a) and (b). We made the following changes.

Diagram

Description automatically generated

1. *Section 3 needs structural editing: currently it states that there are “five aspects to consider…” before enumerating six questions and elaborating in five subsections (plus summary). I think that item 5 in the list of questions could be merged with item 4 to reconcile this. Also, the summary suggests that users adopt the minimum of 4 different k limits, is there one missing?*

*“…item 5 in the list of questions could be merged with item 4 to reconcile this.*” – Indeed. The original item 4 and item 5 basically express the same idea. They are merged and rewritten as the current item 4.

“From which *k* is the variability in the *Ak* and *ϕk* time-series dominated by the measurement uncertainty of *Ak*’s and *ϕk*’s?”

“*Also, the summary suggests that users adopt the minimum of 4 different k limits, is there one missing?*” – Nothing is missing as item 2 “What is the distribution of errors in *Ak* and *ϕk* due to photon noise?” is not a constraint for *k*.

1. *(also) Section 3: It would be useful for illustrative purposes to state what k limits are obtained for the analyses done in this study. Maybe this could be a table?*

We added in Section 3.6 “, which is min(20,20,11,12) = 11 for the HARPS-N three years solar data (Section 5)”.

1. *Section 3.2: The first two sentences refer to “linear projection of the spectra” and “Gaussian noise in the spectra” where it should read “CCF” for “spectra.” The CCFs themselves are already transformations of the original spectra, so the assumption of Gaussian noise in the CCF needs to be explicitly justified.*

Yes, it should be CCF. CCF is the linear transform (sum) of the spectrum. Gaussian noise in the spectrum naturally means Gaussian noise in the CCF, which we think is not an assumption but a generally accepted saying.



1. *Section 3.5: This limit doesn’t seem quite right… shouldn’t the number of spectral lines used to construct the CCF also factor into the velocity resolution? A HARPS CCF clearly has meaningful & coherent structure on length scales smaller than 2.5 km/s, or else it would be impossible to measure RVs at the precision HARPS achieves.*

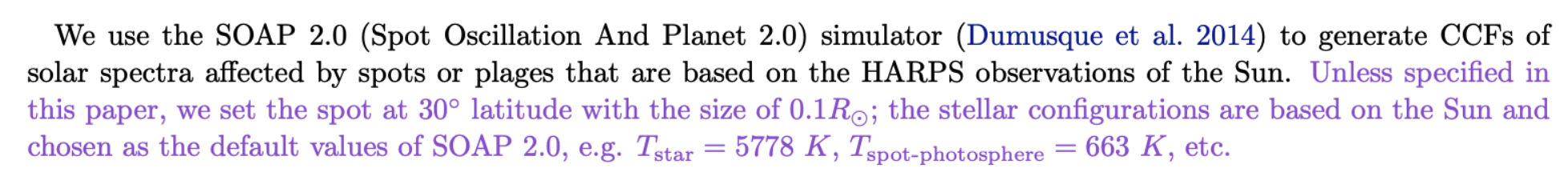
This subsection talks about the effect of instrumental resolution on parametrising CCFs. To avoid misunderstanding, we emphasise “instrumental resolution” wherever resolution appears.

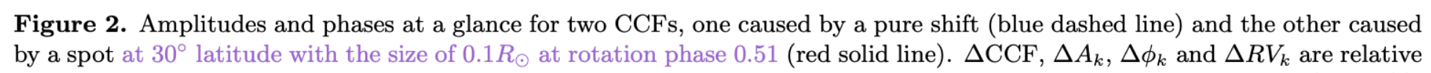
Text

Description automatically generated with medium confidence

1. *Table 1 is not necessary as a table; it could be stated in the text or as a footnote.*

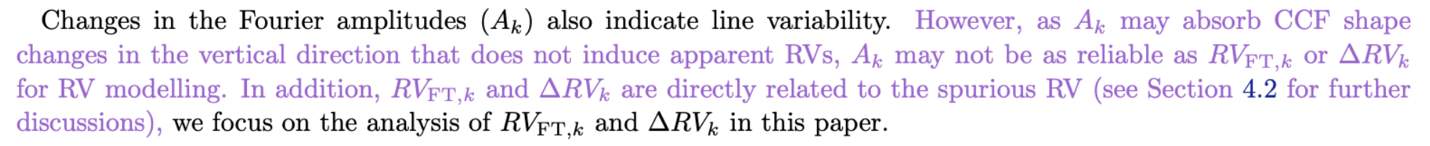
Changed as suggested in Section 4. The other places (Figure 2, 3) that referenced Table 1 have also been updated accordingly.





1. *Figure 2: I would interpret the top-right panel as saying that at small k, the Fourier amplitude is more informative than the phase in disentangling the spot-driven signal from the true RV shift, however the amplitude information is not employed in most of the following analysis. Can this choice be discussed/justified more in the text?*

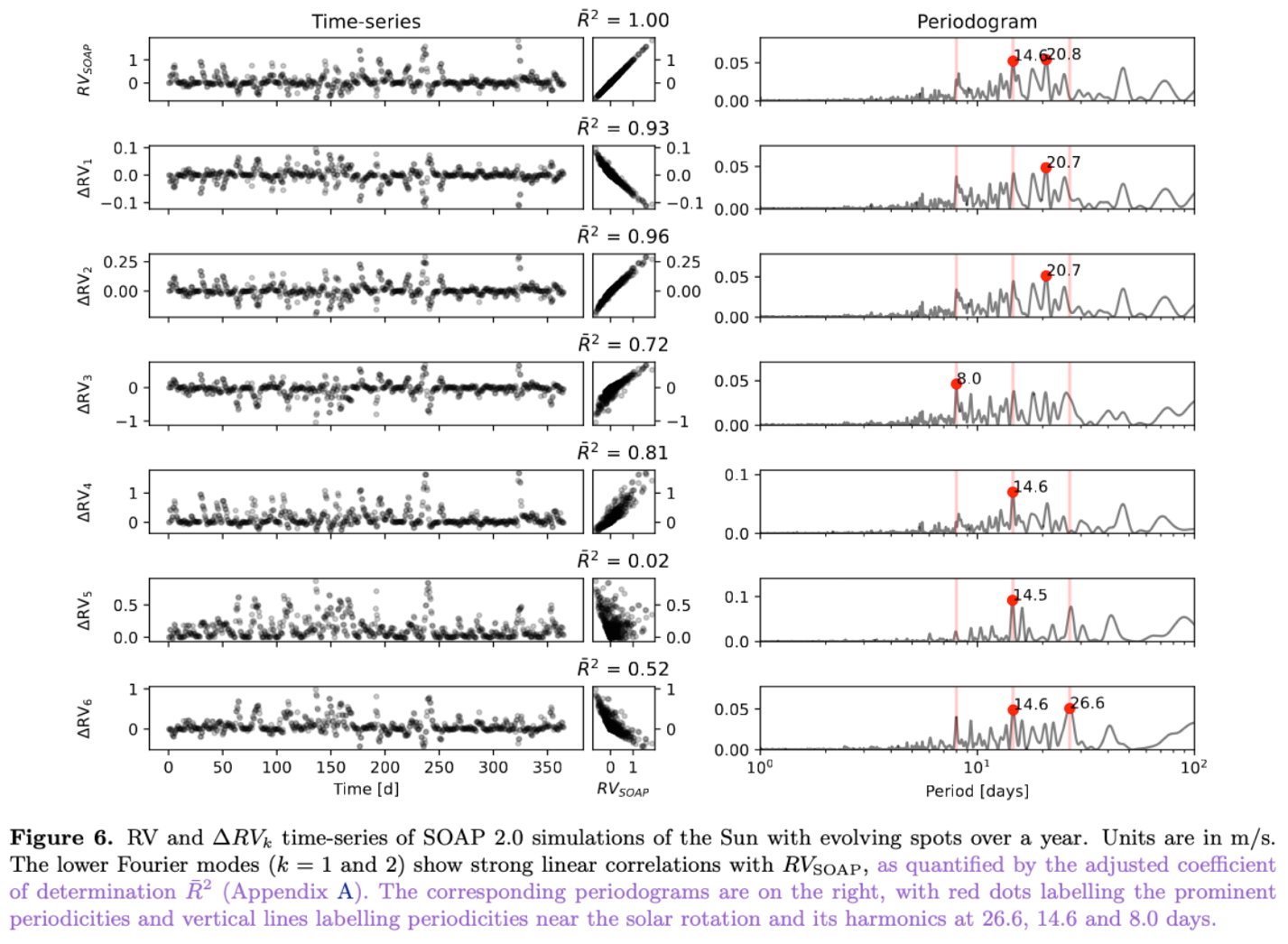
We justify why ∆*RVk* is preferred over *Ak* in Section 2.5.



As for “*the top-right panel as saying that at small k, the Fourier amplitude is more informative than the phase in disentangling the spot-driven signal from the true RV shift*”, we think although the visual effect of ∆*Ak* might mislead the reader to think so, what matters is whether the measurements of *Ak* and *RVFT,k*for a line deformation and a line shift are significantly different, and this significance is not comparable without measurement errors, which is not simulated in the demonstration in Figure 2.

1. *Section 4.4: This section and Figure 6 give really nice context for the following analysis. It may be worth showing or at least commenting on the periodograms of the various k components, for maximum comparability to Figures 7 & 8. I think it would actually be worth going through the entire analysis on this simulated data, including the PCA and subsequent filtering steps; if the intuition is correct that reducing dimensionality with PCA is motivated by the limited number of mechanisms changing the CCF, then the simulated data with only spots should have fewer high-scoring PCA components and it should be entirely contained in the S-PC short-term variable components, right? Is that the case?*

We had actually inspected the periodograms of the SOAP RVs and the ∆*RVk* but they were not very interesting in this ideal simulation and so were not included in the original manuscript. However, considering the readers might ask the same question, we attach the periodogram for their reference.



If we were to continue with PCA for ∆*RVk*, based on the high correlations of the two modes that we see in Figure 6, the first (i.e., most prominent) PC score would be similar to the RV time-series as well as the first two modes, which would account for over 96% of the variance of the data. The rest principal components explain less than 4% of the variance.

As a result, “*the simulated data with only spots should have fewer high-scoring PCA components*” – Yes, and only one PC is needed to explain over 96% of the variance.

“*it should be entirely contained in the S-PC short-term variable components*” – Ture, but this is only because the SOAP simulated spectra do not exhibit long-term variability, such as magnetic cycle and instrumental instability that we see in the HARPS-N observations.

We also want to emphasise that while the FIESTA approach is standardised, the approach that we work on the HARPS-N data in Section 5, (i.e., PCA on ∆*RVk* 🡪 separating out the long-term and short-term variabilities 🡪 multiple linear regression modelling) does not necessarily need to be a standard procedure. We do not want to leave the reader the impression that this is the only way we make use of the FIESTA outputs, and in the case of Section 4.4, there is no need to proceed with the subsequent analysis as reasons mentioned above.

1. *Related to the above point: can you test for overfitting by injecting & recovering a true Doppler shift? I assume that doing the linear regression modeling from Section 5.5 ought to remove the spot effects while preserving true Doppler signals, so can you show that is actually the case?*

We agree with the reviewer that injecting a true Doppler shift is crucial in testing if we are overfitting the RV variabilities. It requires simultaneous modelling of the line variability RV and the planetary Keplerian orbit. As a robust test, multiple amplitudes and periods for the injected planet need to be evaluated. Considering the length and scope of the current manuscript, we prefer not to include the analysis for the moment but use the Lasso regression to address overfitting (e.g. Section 5.5, 5.6). Nevertheless, planet injection is definitely the next phase of exploration.

1. *Figures 7 & 8: These figures might be easier to interpret if you add semi-transparent shaded regions in different colors around the significant periodicities enumerated in Section 5.2, so the viewer can see at a glance whether different Fourier components correspond in part to different mechanisms.*

Figures 7 and 8 (as well as 6) are updated as suggested.

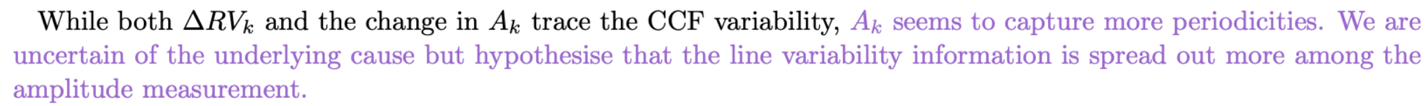
Table

Description automatically generatedTable

Description automatically generated

1. *Section 5.2: I am not sure I understand the point about the A\_k periodograms being noisier because the A\_k values are affected by continuum normalization, can you elaborate? Or potentially address this with similar plots of A\_k for the simulated data?*

We think the term “noisier” is vague and have rephrased the sentence as



Below we attach the *Ak* periodogram of the simulated data for the reviewer’s reference (not presented in the manuscript). It does capture more periodicities, but these extra periodicities do not have straightforward meanings, which we originally referred to as “noise”.

A picture containing table

Description automatically generated

1. *Section 5.3: The PCA step of analysis needs to be motivated a bit more. The reasoning is that certain aspects of the line deformation might map into multiple Fourier components, so PCA could separate them out, but then the PCA components do not in fact show an obvious interpretation (e.g. “PCA component #1 looks like rotational modulation”). So why continue on with the PCA parameterization in Sections 5.4 and 5.5, instead of returning to the chief Fourier components and analyzing those? If neither option is particularly interpretable, it seems generally better to stay closer to the data rather than adding in an extra transformation.*

We believe that the PCA approach helps with the interpretation of the Fourier modes by preserving most of the periodicities with only 3 principal components instead of 20 Fourier modes. If we were to use the first few modes for the line profile parametrisation, the PCA approach would contain more information than the same number of Fourier modes. We also discussed in Section 7.2.1 that future regularisation of the PCA scores will potentially separate out the physical processes of different timescales. When it comes to multicollinearity in the regression models, dimension reduction is normally required to pick up the most significant variables and avoid overfitting. In this manuscript, we have used both PCA and lasso regression.

We have rephrased the text and added following in Section 5.3:

“For this reason, there could be redundant features or redundant information if we were to use all the 20 modes in ∆*RVk* as features to parametrise the line profile variability. In addition, as we will see in Section 5.5 where we will fit the RV data in a regression model, dimension reduction approaches such as PCA and regularisation are normally required to pick up the most significant variables and avoid overfitting. PCA projects the higher dimensional data onto the lower new principal axes, which become the new coordinates known as the PCA scores. We can use the PCA scores in a lower dimension for describing the data, albeit neglecting an residual error term.”

In addition, we added the next paragraph to justify the choice of *k*max for PCA.

Text

Description automatically generated

1. *(also) Section 5.3: Why do PCA on Delta RV\_k, as opposed to something like Delta RV\_k multiplied by A\_k (which might naturally resolve the need to downweight the higher frequency modes), or both Delta RV\_k and Delta A\_k?*

The answer can be implied in Section 4.2 – “Weighted mean of *RVFT, k*” where we show the weighted average of *RVFT, k* is the same as the measured apparent RV, which is saying the apparent RV is a linear combination of *RVFT, k*. Similarly, it would make more sense to say the line variability RV is a linear combination of ∆*RVk*. *Ak*, on the other hand, can be treated as the weights.

1. *Section 5.4: The filtering step here could benefit from a little bit more explanation/discussion, since at this point in the paper there are a LOT of different “frequencies” in play!*

We hope the following changes can better explain what we do with the kernel filtering.

Text

Description automatically generated

For clarity, all “Fourier frequency modes” or “frequency modes” are rewritten as “Fourier modes” in the manuscript.

1. *Section 5.5: The justification of the regularization amplitude choice is a bit too vague. Could you use a cross-validation approach?*

Thanks for the suggestion. We have adopted a cross-validation approach in determining the regularisation amplitude .

(Section 5.5)

Text

Description automatically generated

(Section 5.6)

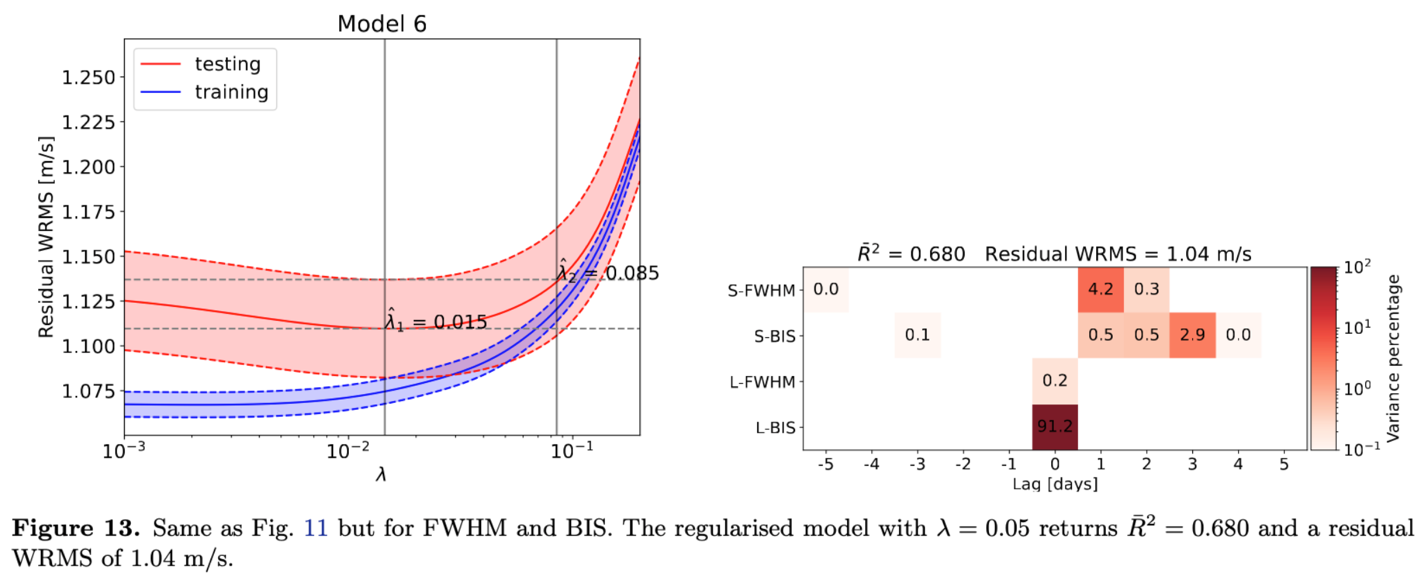
Text

Description automatically generated

Chart, histogram

Description automatically generated

(Section 5.7)



Here are the WRMS - plots for the other Models 1, 2, 4 and 5 for the reviewer’s reference. They are mentioned in the texts but not shown in the paper.



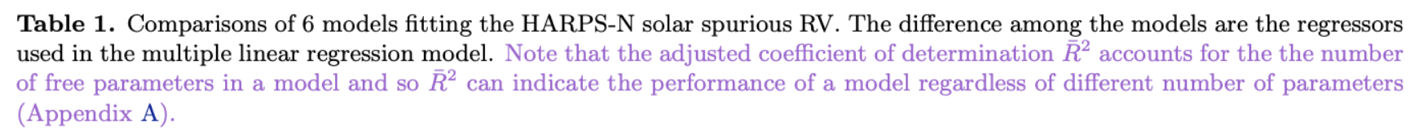


1. *Figures 11 & 13: The colorbar needs a label, and would be better remapped to log scaling to show detail.*

They are both updated as suggested (see the corresponding figures in the previous reply).

1. *Section 5.6: Can you demonstrate that the lagged multiple linear regression improves performance beyond the level expected naturally from adding extra free parameters to the model?*

We compared the performance of different models using the adjusted coefficient of determination which accounts for the number of free parameters in a model.



1. *Section 7: The summary of the core FIESTA method is very clearly stated here. I would add to it a concise description of how the FIESTA results may be used to analyze/denoise RV time series, since this aspect is where the manuscript is more difficult to follow.*

We combine the discussion and the summary session into one, which is now Session 7. The detailed approach for analysing the RV time-series is now followed by the summary of the core FIESTA method.

1. *Section 7.2.2: Is the point about the potential use of temporally correlated noise models supported by the results of the lagged regression in Section 5.6? (I think the answer is yes, so it may be worth mentioning here!)*

It surely is. We address the idea as follows in the current 7.3.2:

Text

Description automatically generated

1. *General comment: The manuscript needs language editing for British/American spelling inconsistencies (for example, normalize and normalise are both used).*

normalize --> normalise

modeling --> modelling

1. *Another general comment: it is great to see the Github repository for the analysis linked in the paper!*

Working on it.

1. *We recommend that living code on github repositories place a "frozen" version on Zenodo (or other 3rd party repositories that issue DOIs) and then cite them in the article.*

To be done.

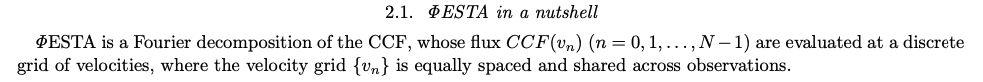
1. *Can equations 1-5 be related to the 'cross-correlation theorem', a generalization of the Wiener-Khinchin theorem? See*[*https://mathworld.wolfram.com/Cross-CorrelationTheorem.html*](https://mathworld.wolfram.com/Cross-CorrelationTheorem.html)*and*[*http://www.ee.ic.ac.uk/hp/staff/dmb/courses/E1Fourier/00800\_Correlation.pdf*](http://www.ee.ic.ac.uk/hp/staff/dmb/courses/E1Fourier/00800_Correlation.pdf)*.*

Equations 1-5 are definitions of discrete Fourier transform, its inverse form and the related quantities. The form can be replaced by any one-dimensional array input and it does not depend on how the spectral CCF is derived.

We assume the reviewer wanted to say can be expressed as where is the cross-correlation symbol, is the stellar spectrum and is the line list template used to build the CCF. The Fourier transform of becomes , i.e., proportional to the dot product of the Fourier transform of the stellar spectrum and the Fourier transform of the line list template. Now that the phase of stellar spectrum and the line list template in the Fourier domain adds up linearly, the phase changes in the Fourier transform of CCF is directly linked to that of the stellar spectrum as the line list template is considered static, i.e., a frequency dependent phase shift in the stellar spectrum results in the same phase shift in the Fourier transform of the CCF. It is an interesting conclusion derived from the Wiener-Khinchin theorem, but we do not use it in the manuscript.

Other changes not mentioned above.

* Added co-author Chris Tinney.
* The description of the CCF velocity grid is made clear and concise at the beginning of Section 2.1.



­