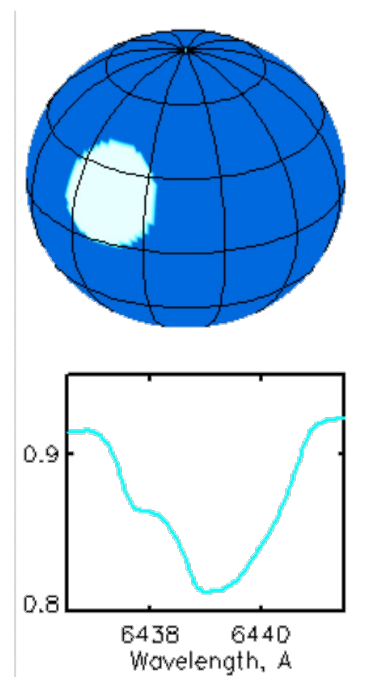
Fundamental limit to RV precision taking into account the stellar activity

Joe Ninan[1], HPF and NEID Team

[1]: Tata Institute of Fundamanetal Research, Mumbai, India

Abstract:

Can we ever detect Earth like planets around sun like stars using extreme precision radial velocity (EPRV) in light of stellar variability induced jitter?
This is a fundamental question that the EPRV community has been asking. The Carmér-Rao bound on the precision by which one can estimate a parameter from an observation is independent of the method used for estimation. It is completely determined by the sensitivity of the observed quantity to the parameters.
If one has a forward model of stellar variability, that is sufficient to calculate the Fisher Information in the data to obtain the fundamental limit of precision by which we can measure the radial velocity shift hidden in the data.
Here, we preset a simple framework and tool to calculate this limit for any given spectroscopic observation and a forward model of stellar variability.



Do these thoughts keep you awake at night?

- Will we ever be able to disentangle 10cm/s planet signal from stellar jitter in a spectrum?
- Are the stellar variability indicators I measure out of this spectrum sufficient?
- Will my arch nemesis come up with a yet to be discovered stellar variability signal from this spectrum that will help disentangle the planetary signal significantly better?
- Should I spend time developing a machine learning algorithm to discover more complicated higher order stellar activity signal from this spectrum to clean the measured RV data?
- Am I loosing stellar activity information when I am throwing away or averaging out some parts of the spectrum?

Fear not, Information theory has you covered! The information content of any parameter in any measurement is independent of your retrieval method. It can be estimated if you have a forward model that can produce all variations in the measured data.

If we can come up with a forward model (even if the model has no interpretable meaning to its parameters), that is sufficient to calculate the Fisher Information Matrix

This estimate will be only as good as your variability model.

The Fisher Information Matrix

- Let S(θ,v) be the observed spectrum, where θ represents the vector of hidden parameters that parameterizes stellar variability that changes the spectrum, and v represent the radial velocity that Doppler shits the spectrum due to a real planet signal.
- In case of a spectrum where the measurement error σ in each pixel is independent, one can write the Fisher information matrix as summation over all pixels

$$F_{ij}=\sum_{\theta,v}\frac{1}{\sigma^2}\frac{\partial S}{\partial i}\frac{\partial S}{\partial j}+\delta_{ij}\frac{1}{\sigma^2_{prior}}$$

- The matrix inverse of this Fisher Information Matrix will give the covariance matrix. This is the Carmér-Rao bound, and no retrieval method will be able to give a more precise answer than this bound.

Example: PCA based forward model

- Stellar spectrum variability can be represented by a linear combination of Principal component Analysis (PCA) eigenvectors.
- Let P_i be the eigen vectors of stellar variability.

Let S be the spectrum,
and S' be the derivative of S with respect to doppler shift v.
Then,

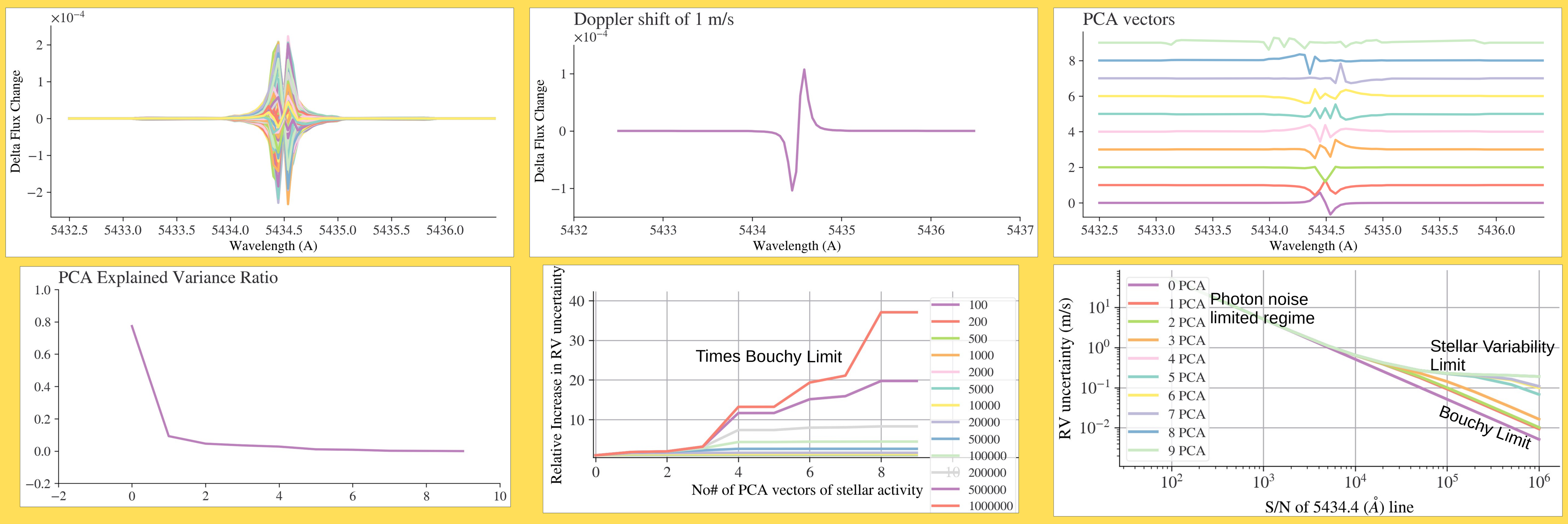
$$S=vS'+\sum_i c_i P_i$$

$$\frac{\partial S}{\partial c_i}=P_i \qquad \frac{\partial S}{\partial v}=S'$$

The Fisher Information Matrix summed over pixels

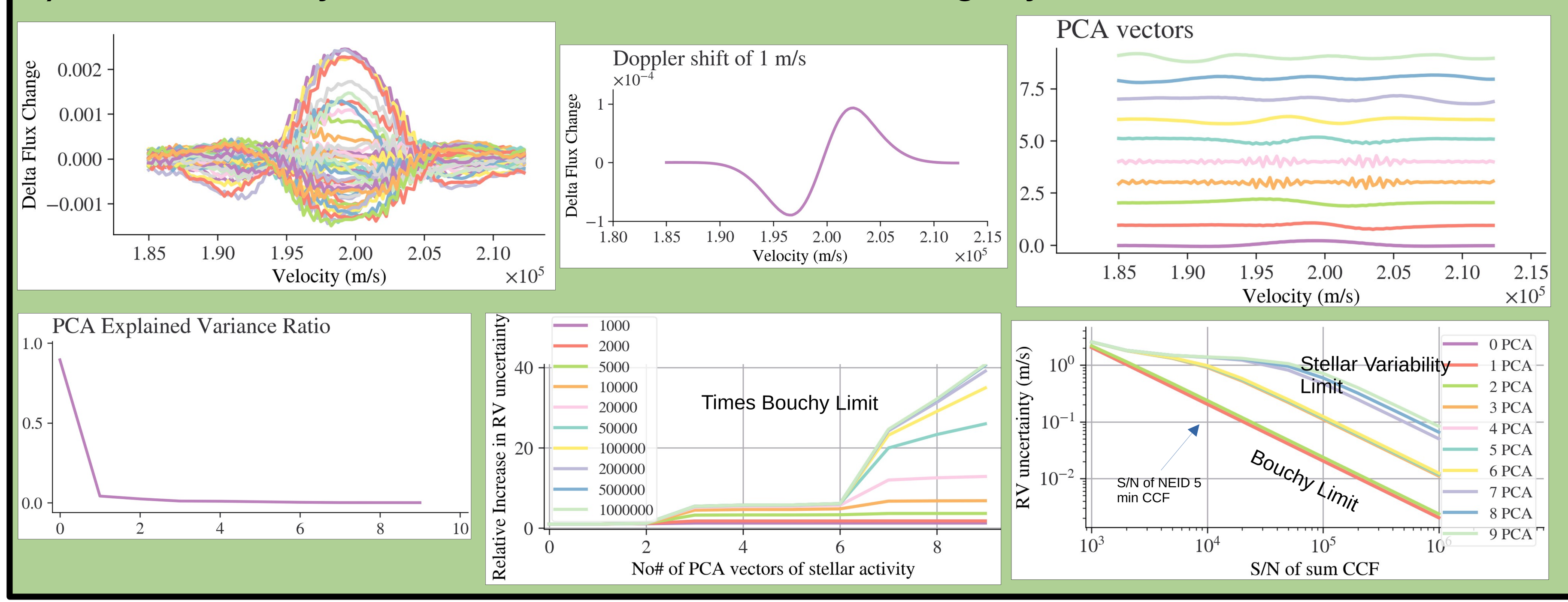
$$F_{jk}=\sum\frac{1}{\sigma^2}\frac{\partial S}{\partial j}\frac{\partial S}{\partial k}+\delta_{jk}\frac{1}{\sigma^2_{c_j\,prior}}$$

1) Impact of granulation on a single line: Based on simulation of 5434.4 A line using GRASS (Palumbo et. al.. 2021)

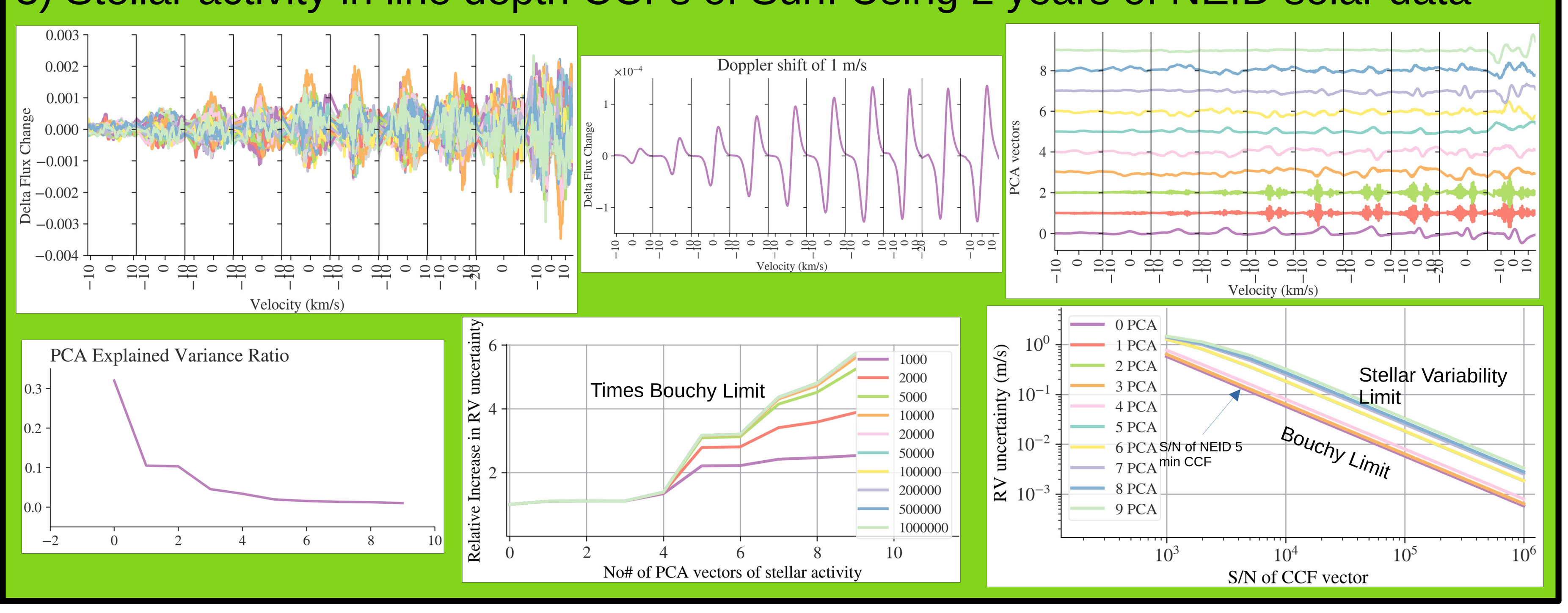


Quantification of stellar activity information loss while averaging lines of different depths for CCF

2) Stellar activity in cumulative CCF of Sun: Using 2 years of NEID solar data



3) Stellar activity in line depth CCFs of Sun: Using 2 years of NEID solar data



Conclusion & Ongoing work:

- Fisher Information calculations are useful to obtain an estimate of Carmér-Rao limit on any single measurement RV precision.
- Our PCA based stellar activity decomposition model shows averaging of lines from different depths to generate a single CCF reduces the best possible RV measurement precision by a factor of ~ 3.5 times.
- PCA based modelling of stellar activity for Fisher information calculation is useful to optimize the preprocessing or region selection of spectra to minimise RV information dilution due to stellar activity.
- Development of Variational Auto Encoder based spectral variability parameterization might give a more compact representation of stellar jitter modelling.
- The current model does not consider the correlated stellar activity information in the time axis of multi epoch spectrum. We plan to extend the forward model to capture this information in time variability using Gaussian Process (Gupta et. al. 2023).

Acknowledgment:
We thank Dept. of Atomic Energy (DAE), Gov. of India and Murthy Trust for funding this research.We thank Prof. Suvrath Mahadevan & Prof. Eric Ford for insightful discussions & suggestions.

joe.ninan@tifr.res.in

EPRV V, Leiden, 2024