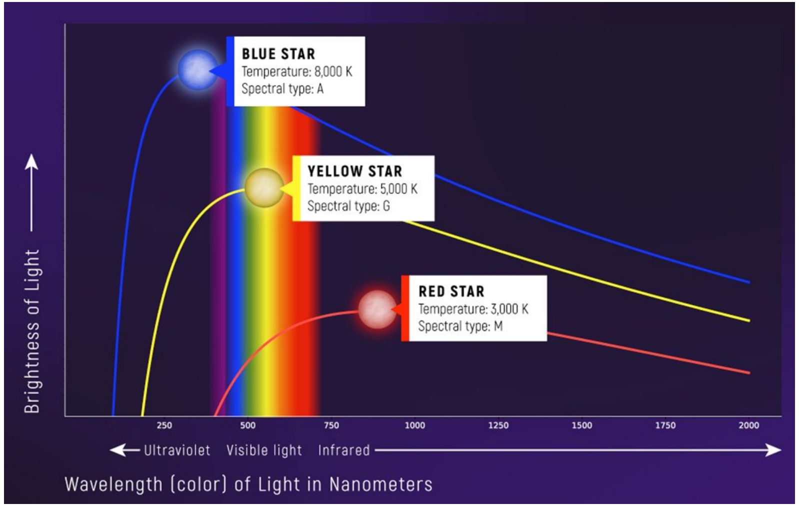
AstroPy Polynomial Fitting to Star Blackbody Spectrums

By James Baldwin, Kevin Barone, Ryan Bennett and Raymond Miller

Motivation

- Background
  - All objects emit light certain wavelengths of light
  - Lower temperature objects (such as living beings) emit primarily in the infrared spectrum
  - Stars emit both in infrared and much more energetic wavelengths, such as UV, X-ray and Gamma rays
- What is blackbody spectra?
  - An idealized relationship between wavelengths of light emitted and their brightness



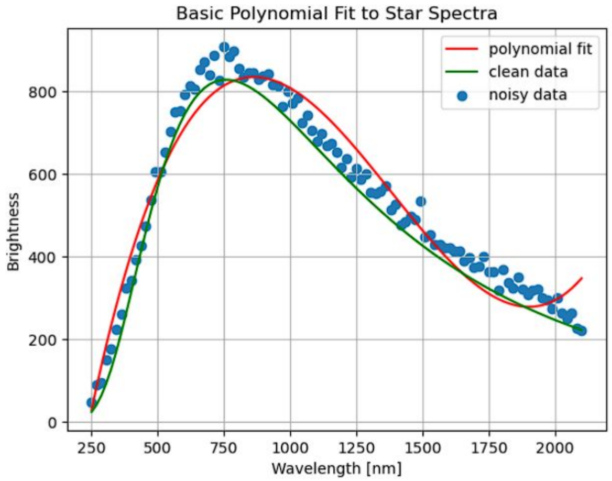
What & How

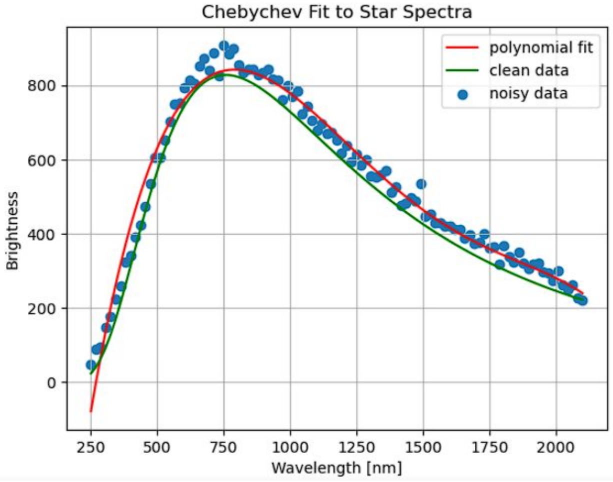
Our Project

- Astronomers need to quickly process and categorize **noisy** star spectra data
- We are testing the **accuracy** and **speed** of different polynomial fittings on this data, using the Python package **AstroPy**
- We will test fits with
  - Basic Polynomial, Chebyshev, Legendre and Hermite basis

What we did

- Create synthetic data for 300 stars, each with a different temperature
  - Run fits
  - Calculate time to create fit
  - Measure performance





Metrics & Quality

- We compared the accuracy and efficiency of four different fitting methods: Legendre, Chebyshev, Hermite and polynomial.

- **Measuring accuracy:** We calculated for the the chi-squared value. The chi-squared value measures how well the fit matches the data. Lower value = better fit. Our function calcChiSq() takes in two arrays. Original, clean data and the fitted, noisy data. We calculated the residuals between the two then divides the residuals by the standard deviation of the fitted data and squares the results.
- **Measuring efficiency:** We calculated the execution time it takes to fit the data using the various fitting methods by capturing the CPU time prior to running the fit and subtracting this from the captured CPU time after running the fit. Then we used these readings to calculate and report the average execution time for each fitting method.

- The results showed that Legendre, Chebyshev and Hermite were more accurate and more efficient than the polynomial fitting.
- The Polynomial fit has a noticeably higher maximum execution time. This suggests that it may be less stable and/or more sensitive to certain inputs.

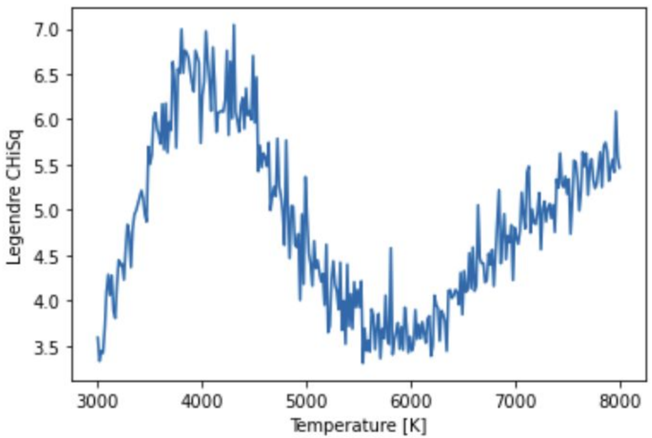
Chi-square Values

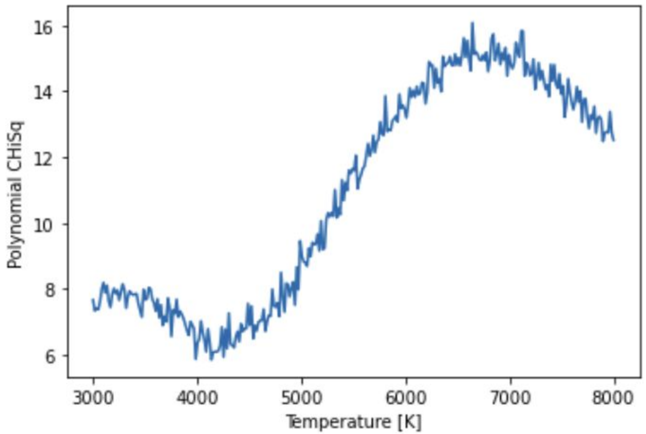
Legendre fit chi-square:
5.4659202048730835

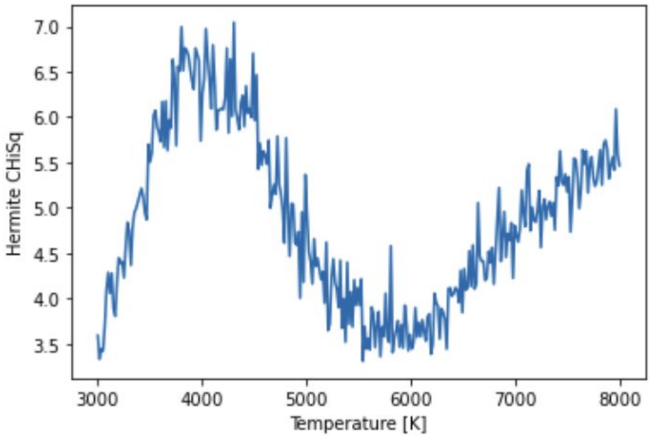
Chebyshev fit chi-square:
5.465920204873062

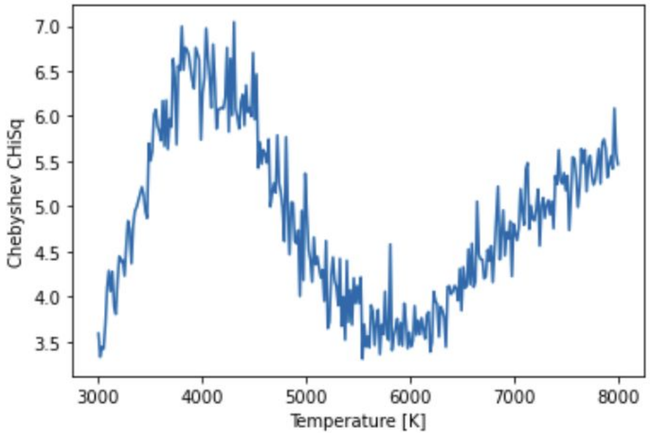
Hermite fit chi-square:
5.465920204873051

Polynomial fit chi-square:
12.519192706357085



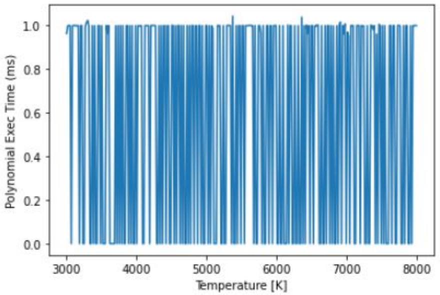


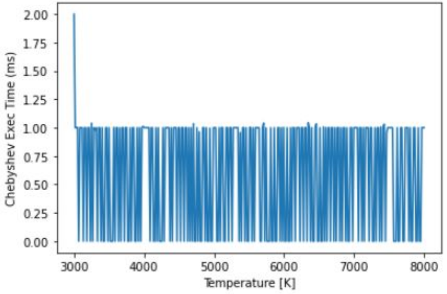


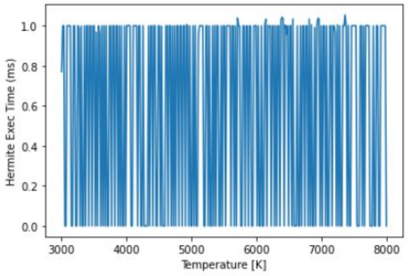


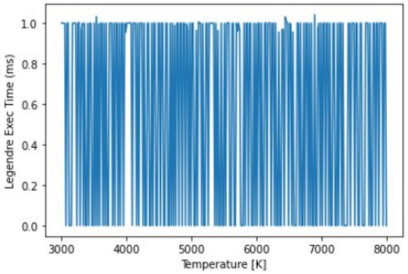
Execution Times of Fits

	Execution Time (milliseconds)	
Fitting Model	**Average**	**Maximum**
Polynomial	0.5841255187988281	4.001140594482422
Chebyshev	0.581789650394275	1.9996166229248047
Legendre	0.5702513000893831	1.0418891906738281
Hermite	0.5778616845013691	1.0535717010498047









Analysis:

Given execution time and performance, the best performing basis are Legendre and Hermite

Impacts

Astrophysicists, cosmologists, and astronomers are constantly working with noisy data sets that can easily reach into the tens of millions. When needing to find a best-fit for this data, a balance between efficiency and accuracy is essential for extracting valuable information in a reasonable amount of time.

POSSIBLE MISUSE

Like many physical scientists astrophysicists, cosmologists, and astronomers can sometimes suffer from certain biases. Sometimes, this means wanting to see a pattern where there is none. While these fitting methods are often accurate in specific situations (ie fitting a blackbody or misaligned spectrum), it is often up to the researcher to determine if such a fit is appropriate to apply to a given set of data.