SYSTEMATICS-DOMINATED SPECTROSCOPIC INFERENCE

Ian Czekala, Harvard → Stanford University
August 23rd, 2016
OVERVIEW

- the fundamental properties of (young) stars
- techniques for measuring stellar properties from spectra (and limitations)
- Gaussian process kernels for spectroscopic "noise" modeling
- emulator for spectral synthesis
- limitations and future challenges
STARS

roughly *single* parameter objects:

models predict how stars evolve with age —

luminosity evolution:

temperature evolution:
MEASURING FUNDAMENTAL STELLAR PROPERTIES

- can be determined for a subset of "special" systems
  - eclipsing binary + RV
  - astrometric binary + RV
  - protoplanetary disk modeling (my focus)

Precise - will be determined for vast majority of stars via spectroscopy
ACCURATE &

For some stars, a major source of uncertainty remains in the accurate determination of \( \cdots \).

Young stars are variable, spotted, and are still accreting.

Move beyond \( \cdots \).

Use **all the data** to determine \( \cdots \).

---

*Constructing a Flexible Likelihood Function for Spectroscopic Inference*

STELLAR SPECTROSCOPY

\[ \text{flux} \quad \text{AND} \quad \text{flux} \]

\[ 40 \, \text{Å} \]
THE PROMISES OF SPECTROSCOPY
AND ITS FAILURE TO LIVE UP TO THEM

a spectrum is a measurement of \( \text{spectra}\), which, in principle, has a \textit{vast} amount of information
e.g., inferring the properties of "fake-data" with S/N = 30 spectra (100,000 pixels!) yields statistical precision of \( \ldots \ldots \ldots \ldots \ldots \ldots \ldots \downarrow \)

\text{No one will ever believe you}
FIRST CONSIDERATION: CALIBRATION
FIRST CONSIDERATION: CALIBRATION

High resolution spectroscopic data is usually acquired from an *echelle* spectrograph, where separate "orders" are later stitched together.

Photon counts need to be converted to units of flux. Continuum-normalization sometimes works, but this is fraught with systematics for cool stars.

Flux calibration of spectra is *hard*, thus, uncertainties need to be modeled.
- include the flux calibration process through nuisance parameters: Chebyshev polynomial coefficients
- inferred along with the stellar parameters
- not necessary to identify continuum regions beforehand
- possible to analytically marginalize over these parameters as well
THE PROBLEM IN A NUTSHELL

stellar spectrum

model function of
1. temperature
2. mass & radius
3. composition
THE PROBLEM IN A NUTSHELL

stellar spectrum

model function of
1. temperature
2. mass & radius
3. composition

residual

careful with
chi squared!
THE PROBLEM IN A NUTSHELL

stellar spectrum

model function of
1. temperature
2. mass & radius
3. composition

residual

careful with chi squared!
AUTOCORRELATION

the "residual spectrum" shows significant autocorrelation power on the scale of a typical line-width
MULTIDIMENSIONAL GAUSSIAN LIKELIHOOD

The structure of $\mathcal{G}$ will be filled in by Gaussian process (GP) covariance kernels.

See also Foreman-Mackey et al. 14, Aigrain et al. 15, Barclay et al. 15
WHITE NOISE

\[
\lambda \quad \text{Å}
\]

-4 -2 0 2 4

\text{noise}\quad \text{residual}

\lambda \quad [\text{Å}]

5,202.0 5,202.5 5,203.0 5,203.5 5,204.0

\text{noise}\quad \text{residual}
GLOBAL (STATIONARY) COVARIANCE

\[ \lambda \] Å

-4 -2 0 2 4

\[ \text{noise} \]

\[ \text{residual} \]

\[ \lambda \text{ [Å]} \]

\[ 5,202.0 5,202.5 5,203.0 5,203.5 5,204.0 \]
LOCAL (NON-STATIONARY) COVARIANCE

outer product of a Gaussian residual described as

noise

residual
use **all the data** to determine

explore posterior with blocked Gibbs sampler
THE NEED FOR "EMULATORS"

synthetic spectra (\( \varphi \)) are expensive to calculate

Multivariate Bayesian Emulator

Habib et al. 07, Heitmann et al. 09
FIND PRINCIPAL COMPONENTS FOR THE PROBLEM
Gaussian processes are convenient to model and analytically marginalize over approximation error

Emulators can provide speed ups of 100x - 1000x
Code base lives online in open-source github repository at https://github.com/iancze/Starfish/

Repository forked by M. Gully-Santiago to include two-temperature component fits, spectroscopic binary fits. All changes in public and trackable.

Gully-Santiago et al. to be submitted
Questions?

DISCUSSION
THE CANNON: DATA-DRIVEN SPECTRA

Ness, Hogg, Casey, Ho, Rix
• Expert knowledge selects sensitive regions
• Mask regions (most of the spectrum) that doesn't fit
• Achieve **highly accurate** stellar parameters
• Is masking satisfying?
FAINT COMPANION SEARCH

single line and double line spectroscopic binary stars (and planet search)

how to determine relative velocity when the template is unknown?

template inference (a GP models part of the search process)
CONCLUSIONS

- spectroscopic modeling for stellar parameters is nearly always dominated by **systematic error**
- flexible **spectroscopic tools** for analysis of young stars are required to move towards realistic uncertainties
- **emulators** make spectroscopic inference possible on a large spectral range
- spectra are fundamental to most astronomical subfields--many opportunities remain for improvement of analysis techniques

Thank you!
FUNDAMENTAL STELLAR PROPERTIES

- young stars are active (accretion, "veiling" continuum)

Herczeg and Hillenbrand 14