Extracting Meaningful Features from Early-Science Radio Data

Matthew Alger (ANU/Data61)
Naomi McClure-Griffiths (ANU)
Cheng Soon Ong (Data61/ANU)

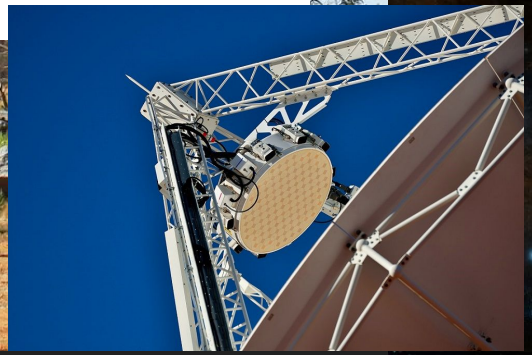
Slides: http://www.mso.anu.edu.au/~alger/aia2019

Australian National University
DATA 61
CSIRO
POSSUM

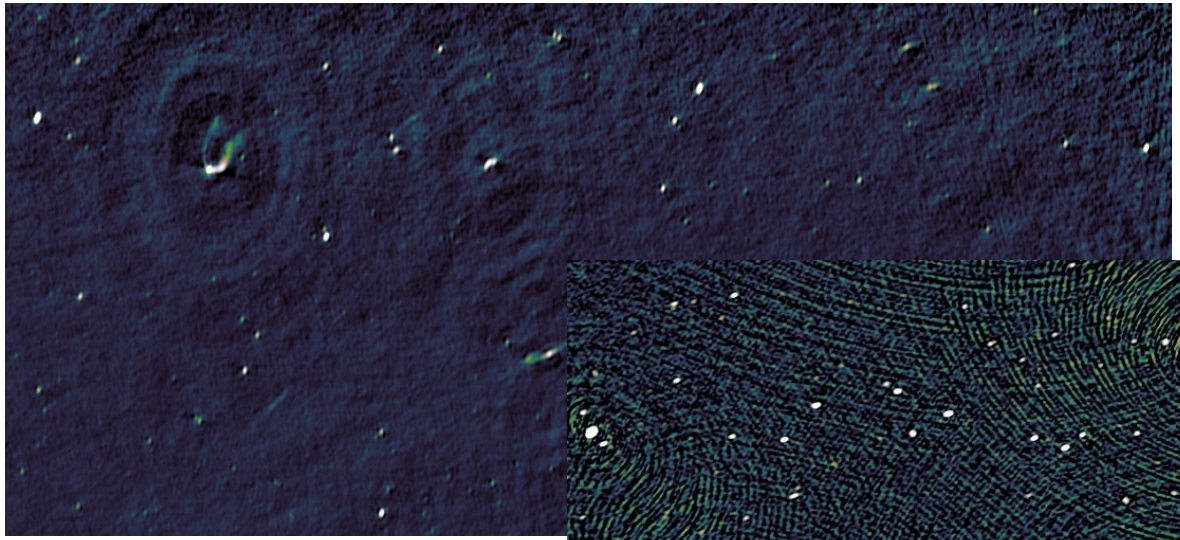
Australian Square Kilometre Array Pathfinder

- Huge 30 deg^2 field of view
- Fast!
- 32 antennae
- >2 PB of science data so far

Image: CSIRO



Early-science data



ELAIS-S1 observed with ATCA (courtesy of Tom Franzen).

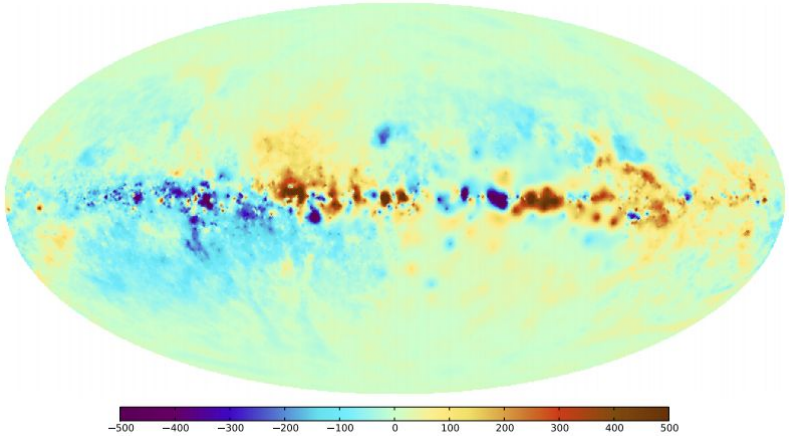
Early POSSUM data observed with ASKAP.

Goals

1. Denoise
2. Get useful features for downstream

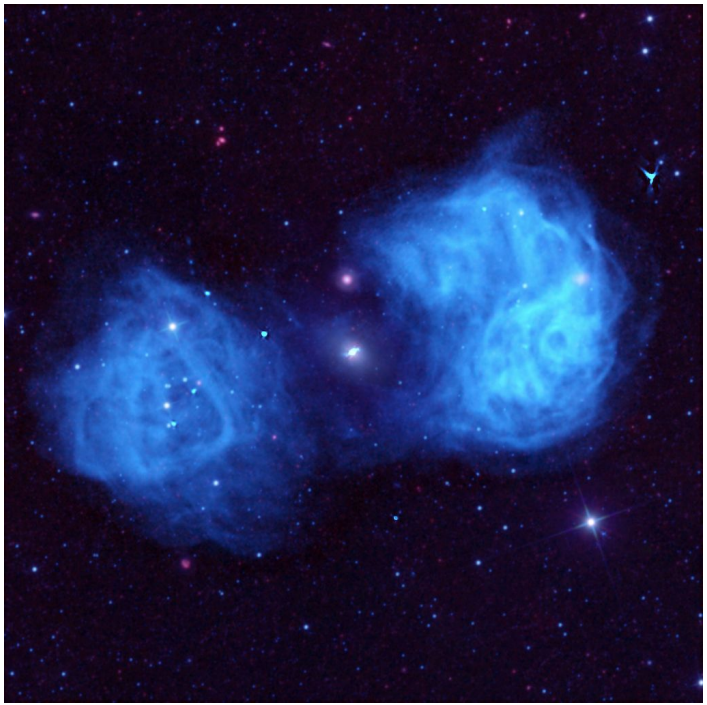
POSSUM

- Polarised “all-sky” survey to complement EMU
- ~1,000,000 polarised radio sources
- Broad benefits to astronomical magnetic field research

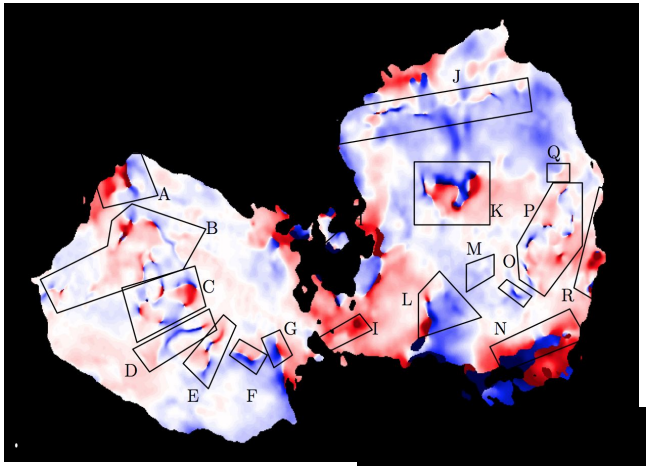


Oppermann+12 Faraday map of the galaxy.

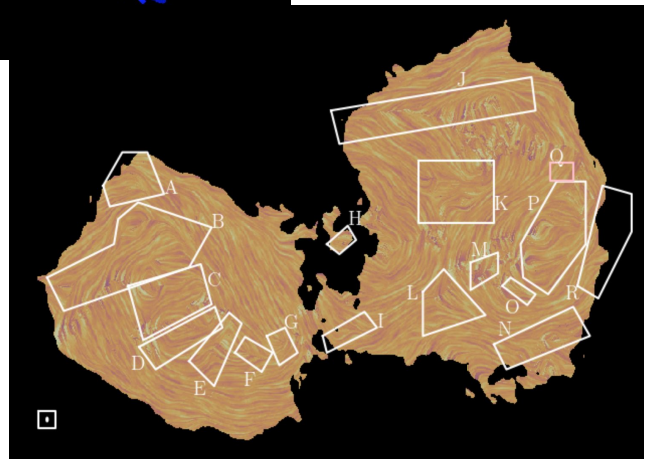
Polarisation



Fornax A in radio continuum (DRAO).

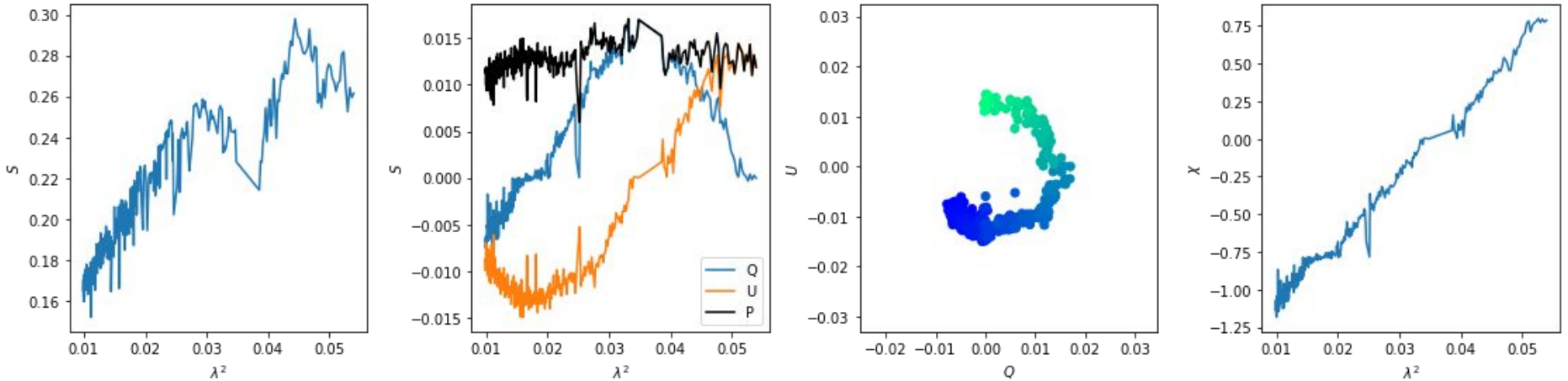


Peak Faraday depth (Anderson+18).



Magnetic field orientation (Anderson+18).

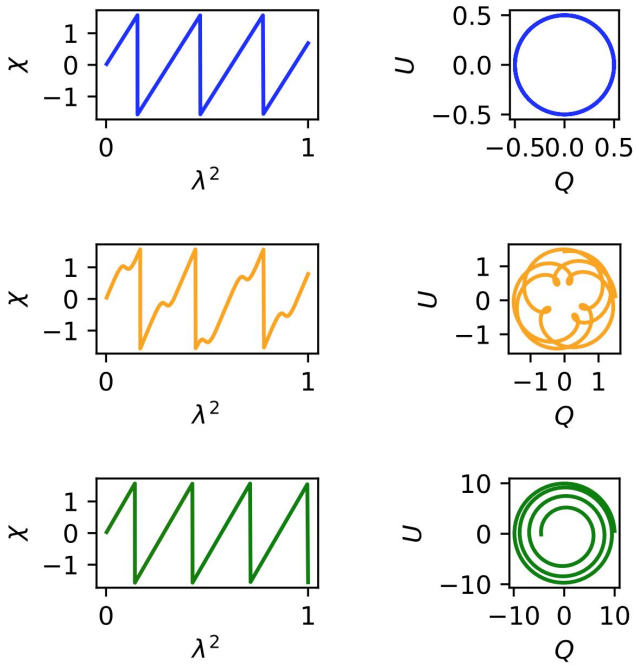
Polarised radio sources



A simple polarised radio source observed with ATCA (courtesy of Jack Livingston).
Left to right: Total intensity, polarised intensity, linear polarisation plane, polarisation angle.

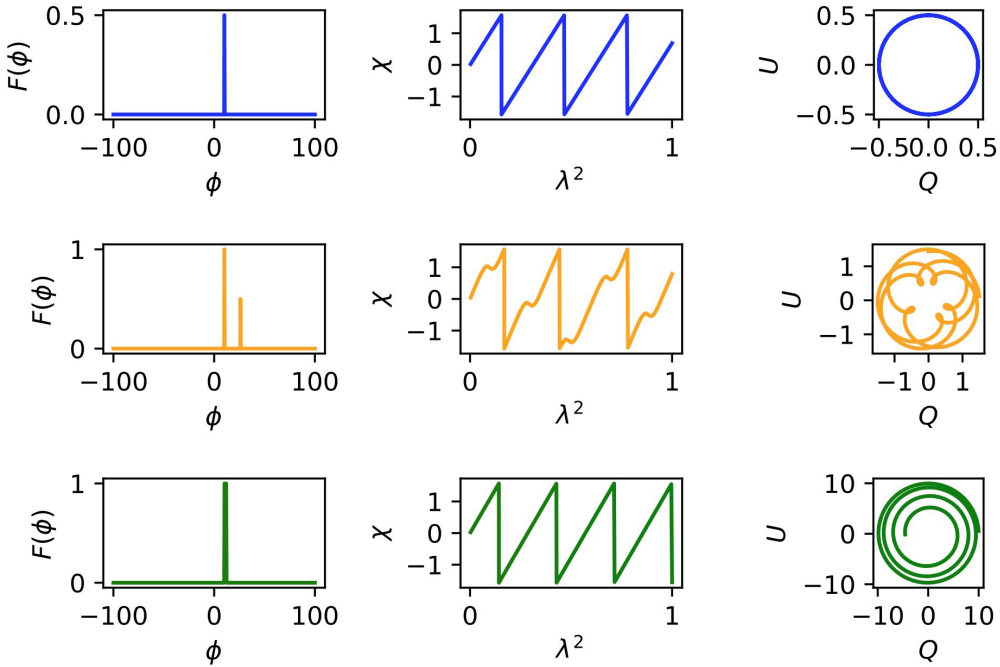
Polarised radio sources

- Simple sources (“screens”) with angle linear in squared wavelength
- Overlapping sources with superimposed rotations
- “Thick” sources with rotation and emission (“slabs”)
  - Depolarisation…



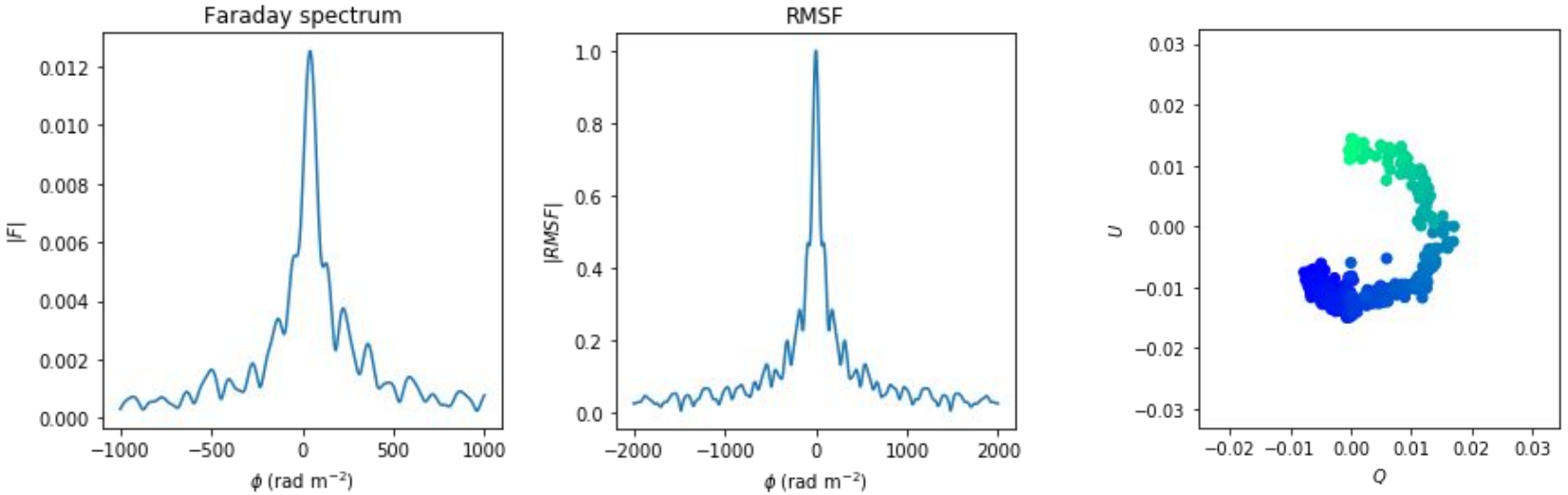
Faraday spectra

- Fourier transform* of polarised spectrum
- Conjugate axis is the *Faraday depth*
- Obvious separation of complexities



Polarised radio sources

Observed spectra noisy and convolved with a spread function (RMSF):



Faraday spectrum of the previous source along with its spread function (courtesy of Jack Livingston).

Feature extraction for Faraday spectra

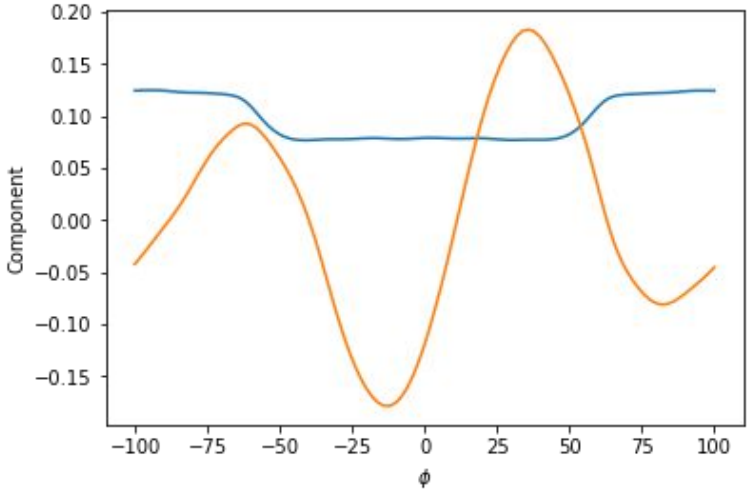
- Good target for polarisation feature extraction
- Extract features from Faraday spectra
- Apply to Faraday complexity classification:
  - Binary classification is nice
  - Well-defined problem

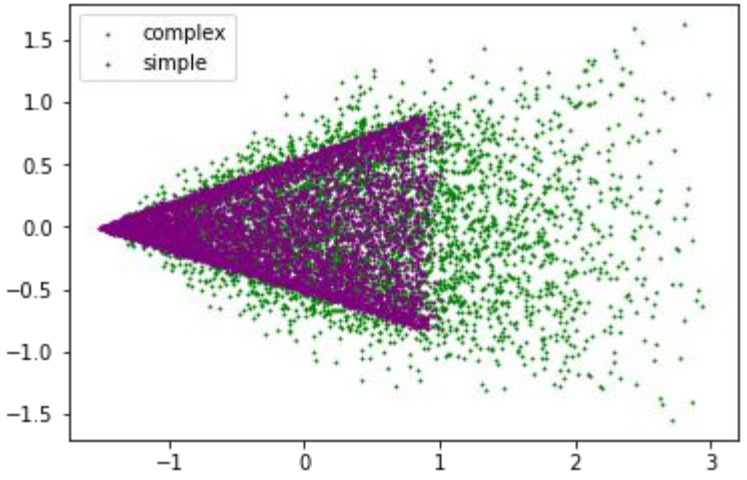
See also...

Classifying Complex Faraday Spectra with Convolutional Neural Networks

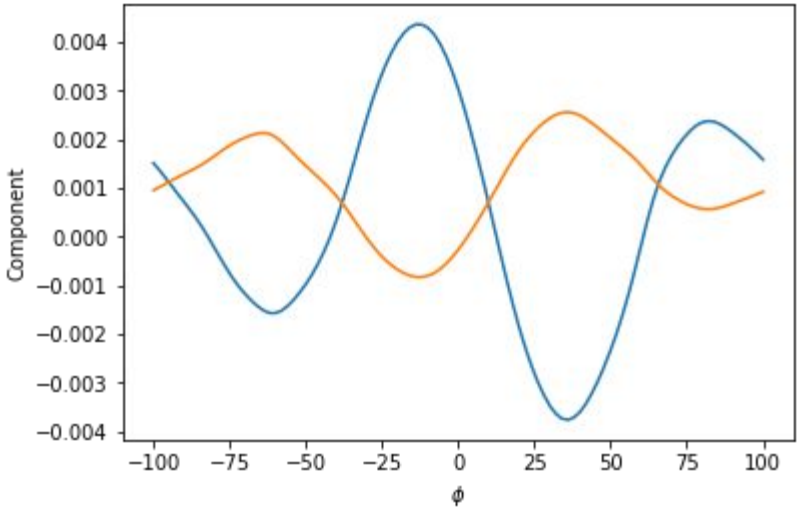
Shea Brown[1⋆], Brandon Bergerud[1], Allison Costa[1], B. M. Gaensler[2], Jacob Isbell[1], Daniel LaRocca[1], Ray Norris[3], Cormac Purcell[4], Lawrence Rudnick[5], Xiaohui Sun[6]

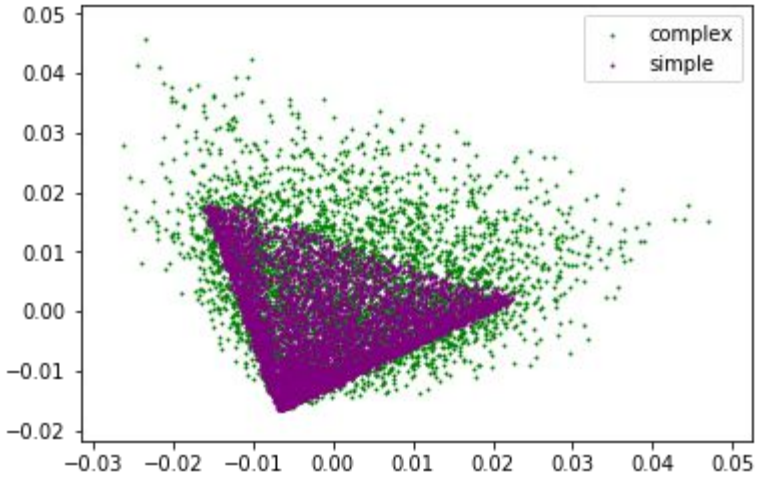
PCA on Faraday spectra





ICA on Faraday spectra

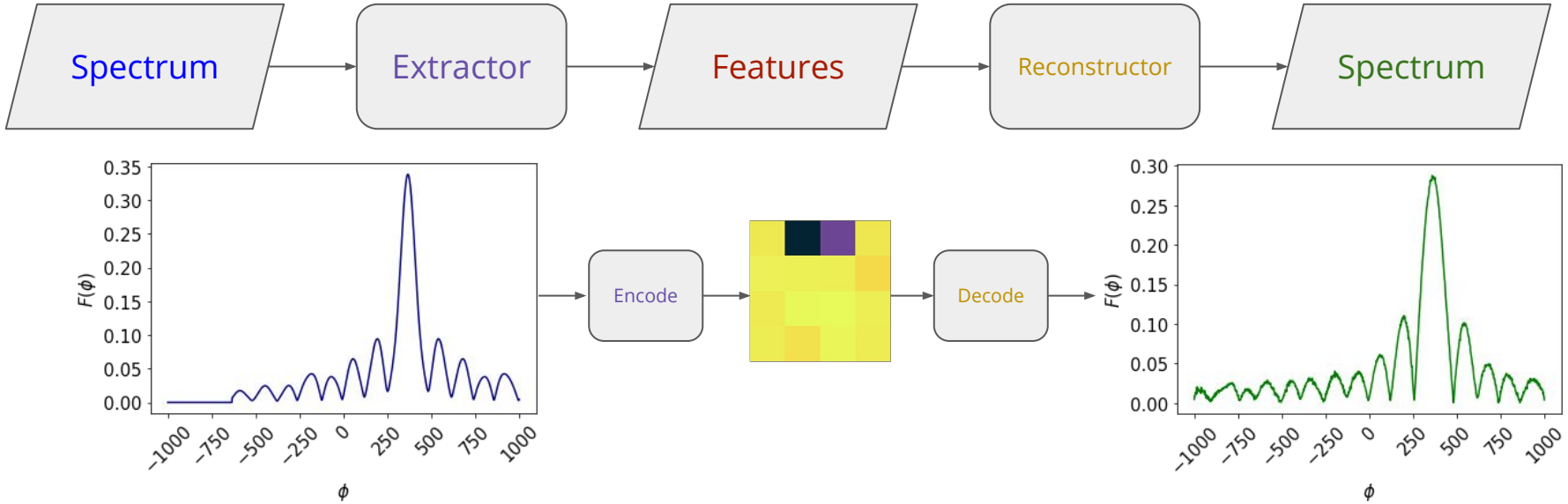




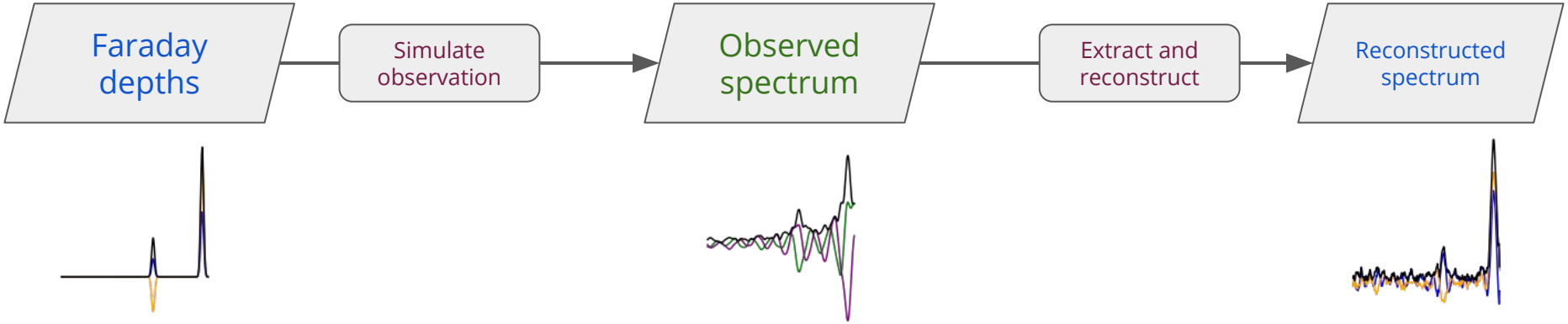
Goals

1. Denoise
    - Add noise, then map back to the no-noise version
    - Learn to “ignore” noise
2. Get useful features for downstream
    - Unclear where these features will best come from
    - Choose a model architecture that explicitly outputs features

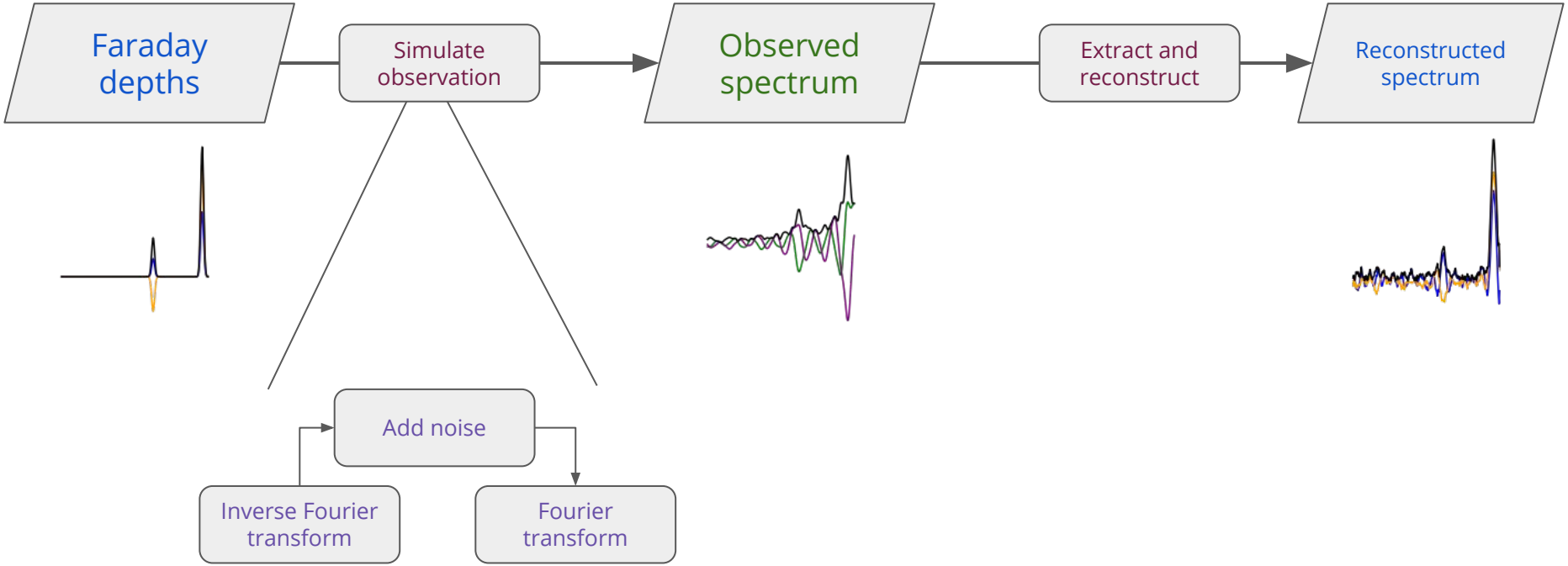
Autoencoders



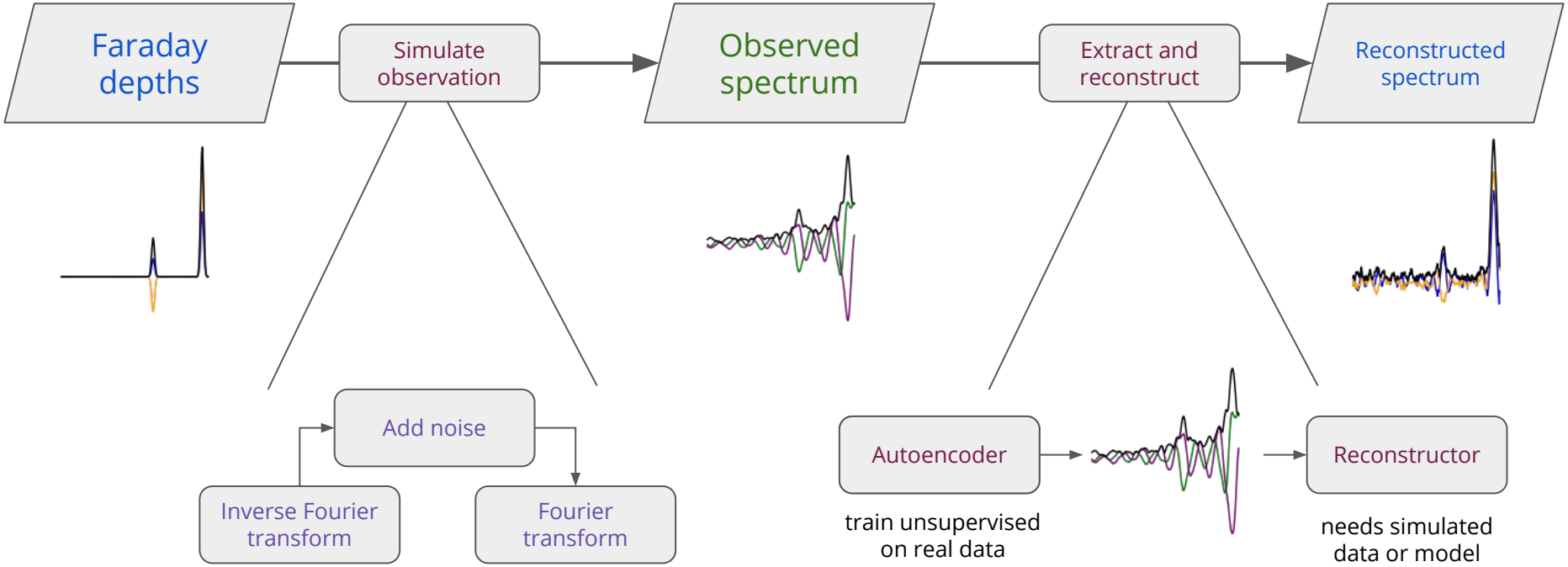
Faraday spectrum feature extractor



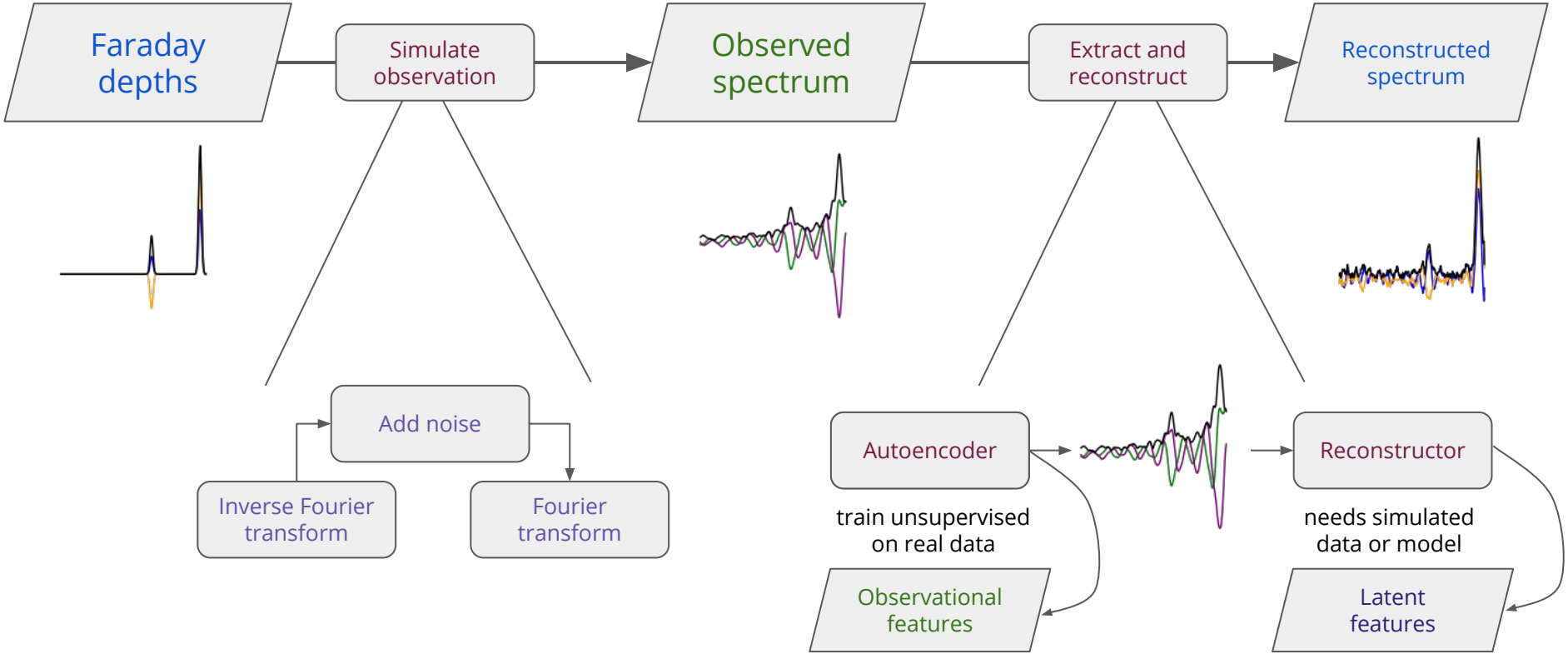
Faraday spectrum feature extractor



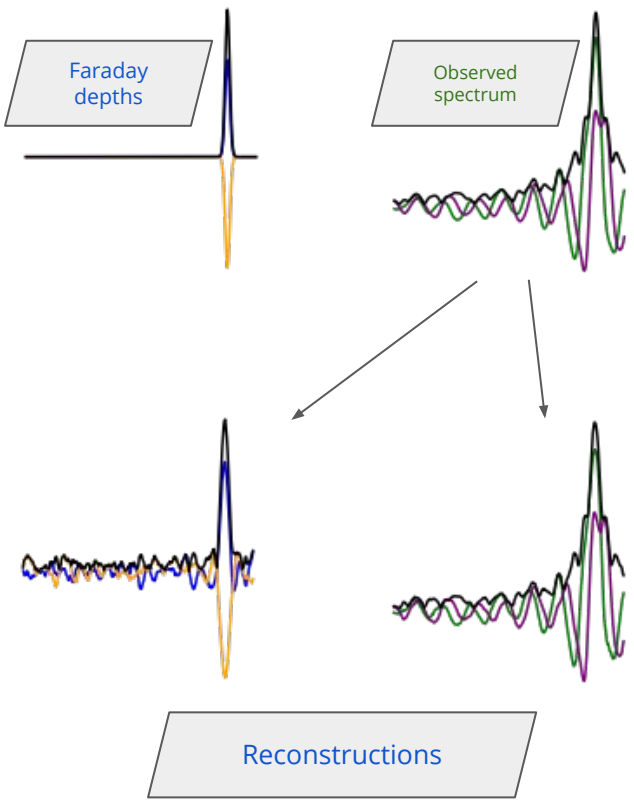
Faraday spectrum feature extractor

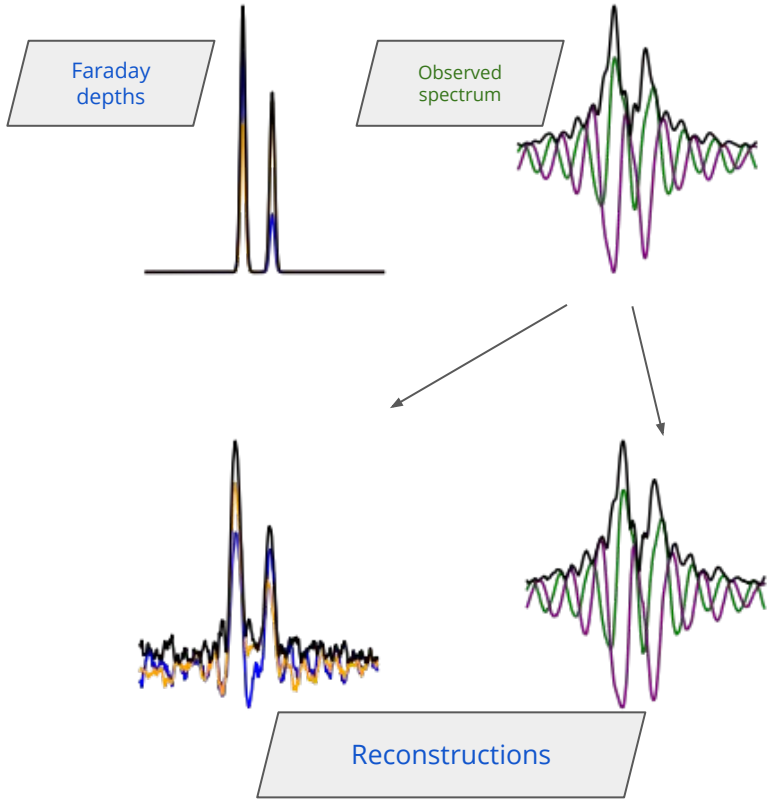


Faraday spectrum feature extractor

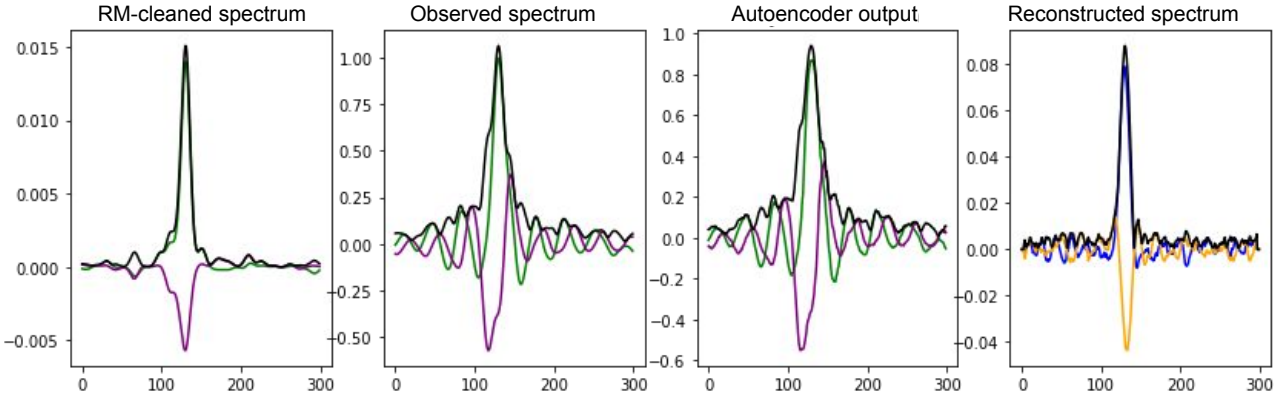


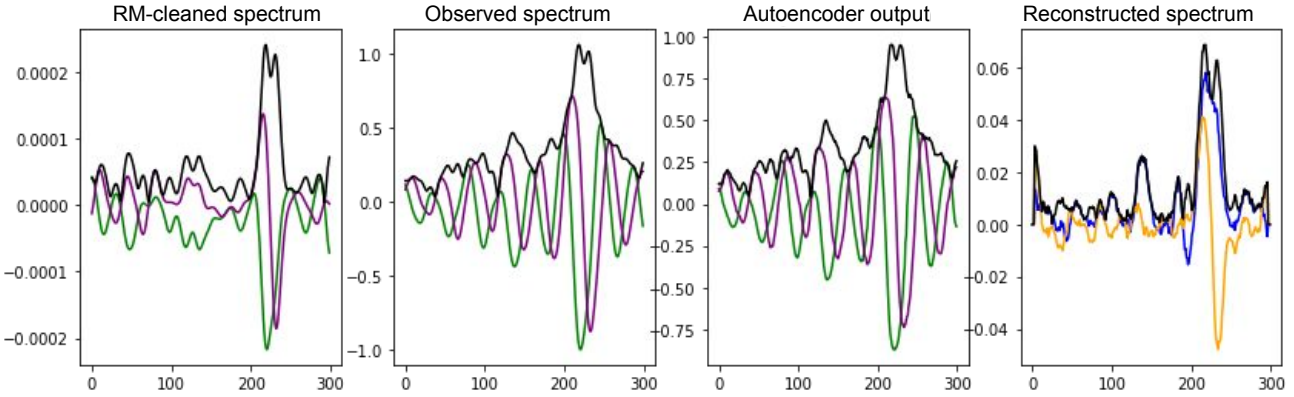
Reconstructing simulated spectra





Reconstructing real data

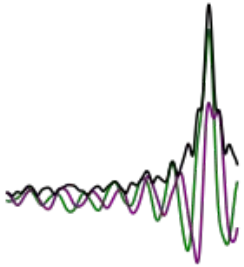


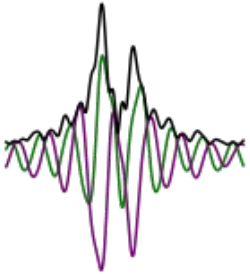


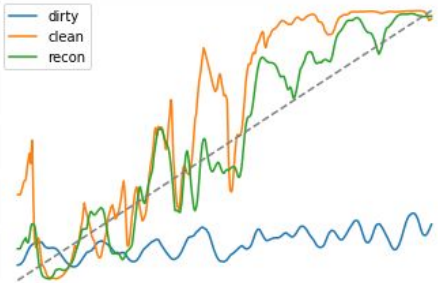
Polarised sources observed with ATCA (courtesy of Jack Livingston).

Using features: Faraday complexity

- Classify spectra as complex or simple
- Not always clear-cut
  - Spectra are noisy
  - Screens can be close together
  - Slabs may look like screens
- Linear models perform alright with latent features





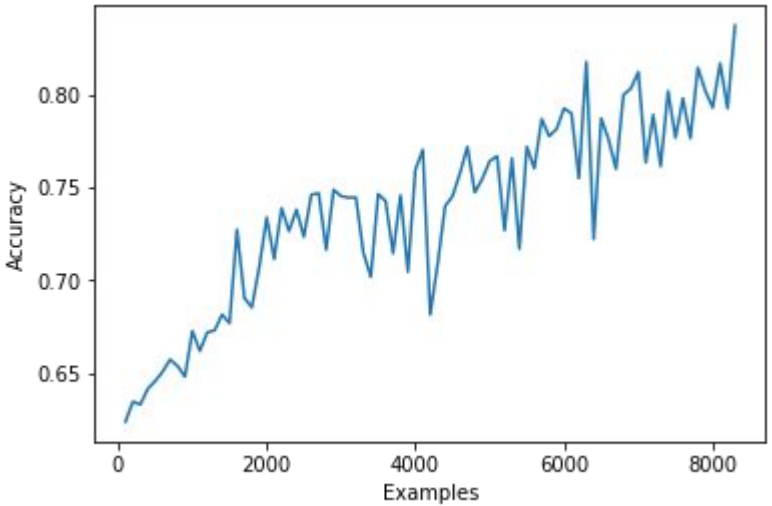


Issues and limitations

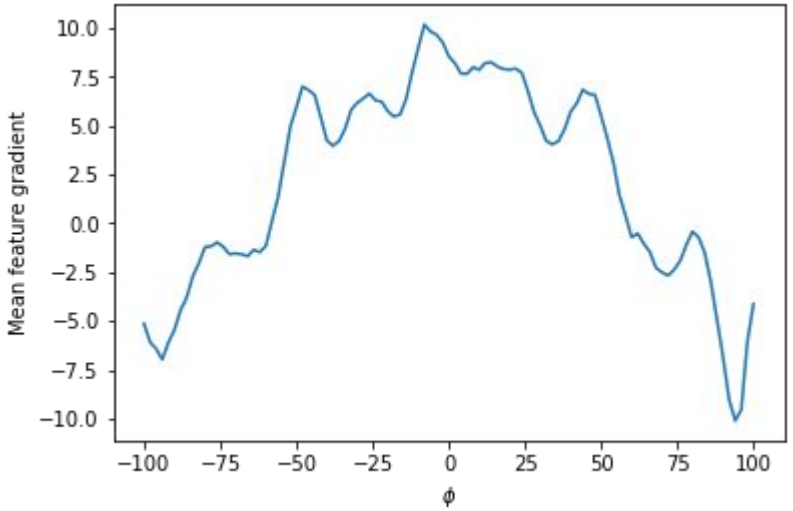
- Extremely sensitive to scale
  - Fortunately scale of *observations* so we can account for this
- Just one RMSF
  - Datasets can have multiple RMSF depending on data quality and issues
  - Different datasets have different RMSFs
  - Still pretty good reconstructions on different datasets!
- Assumes spatially independent spectra

Baseline: Basic convnet

- 1 convolutional layer, 2 dense layers, SGD, 100 epochs
- 85% test set accuracy

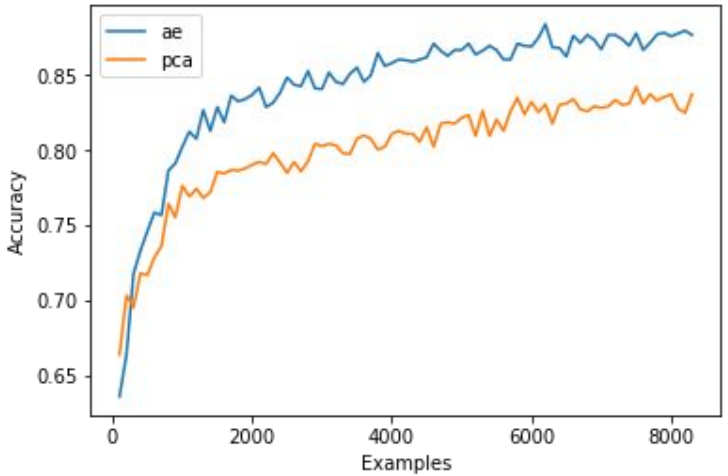


Accuracy as a function of number of examples.

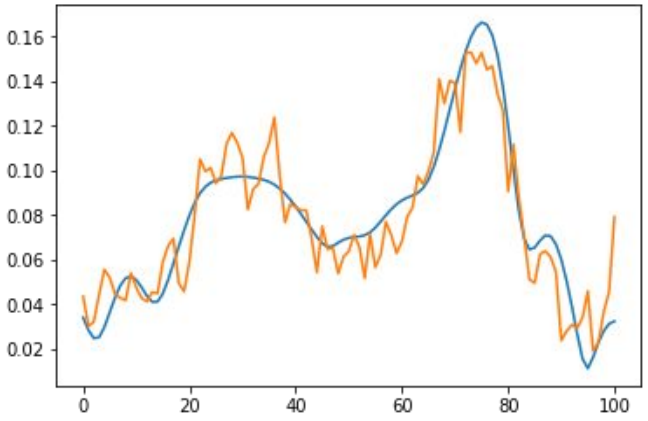


Feature importance for the CNN.

Extracted: 20-feature CAE + RF

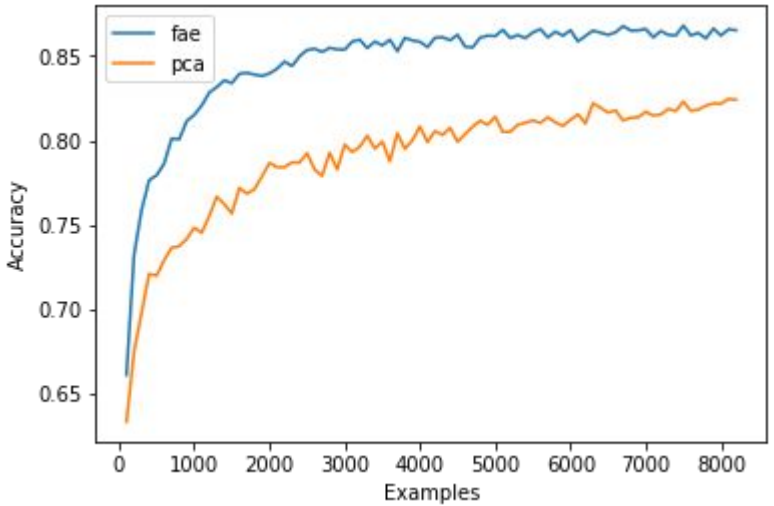


Accuracy as a function of number of examples for AE features and PCA features. Note that we can use the full training set to extract features.

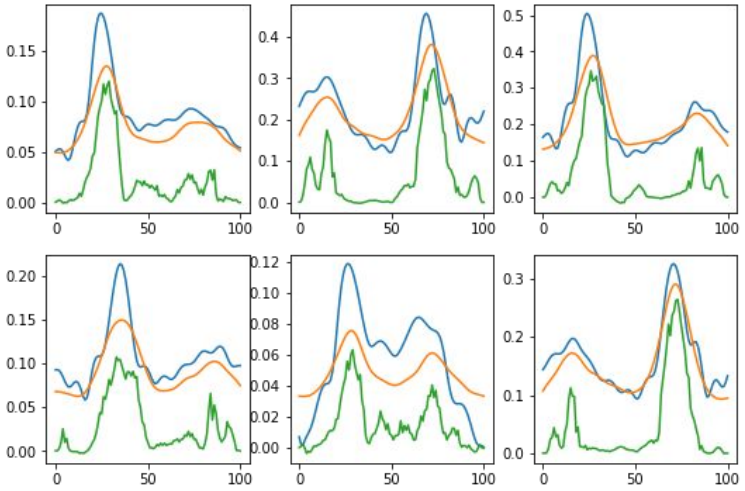


AE reconstruction.

Extracted: 20-feature reconstruction + RF



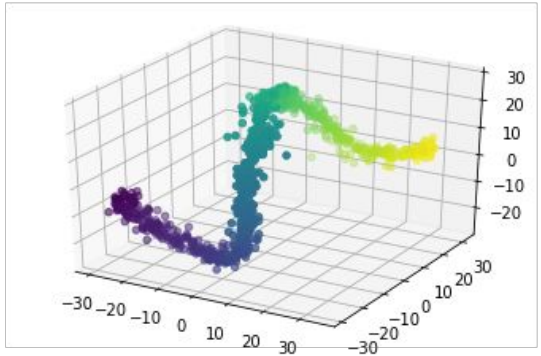
Accuracy as a function of number of examples for reconstruction features (FAE) and PCA features.

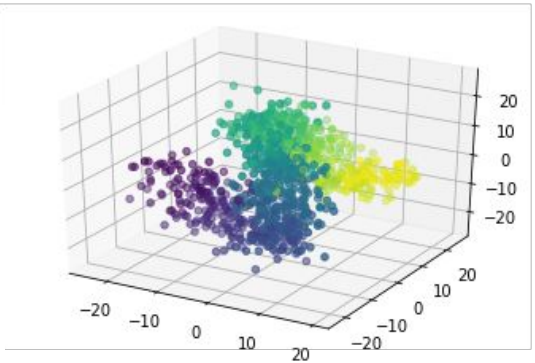


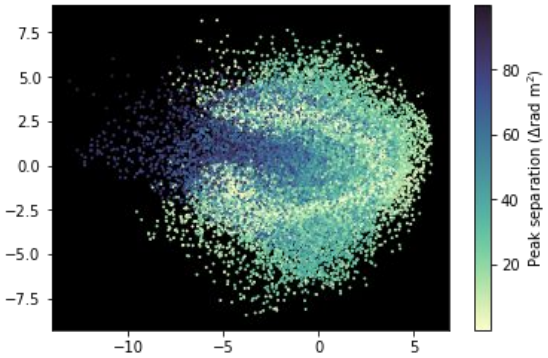
Reconstructions (green) and their convolved counterparts (orange).

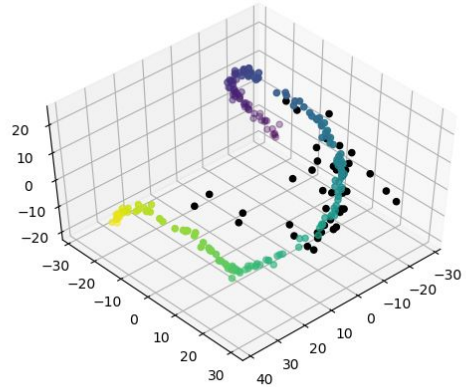
Simple only

- Using a new dimensionality reduction technique, we can extract features from just simple sources
- Simple sources lie on a 1D manifold (plus noise)
- Real observations can be projected onto this manifold to score complexity



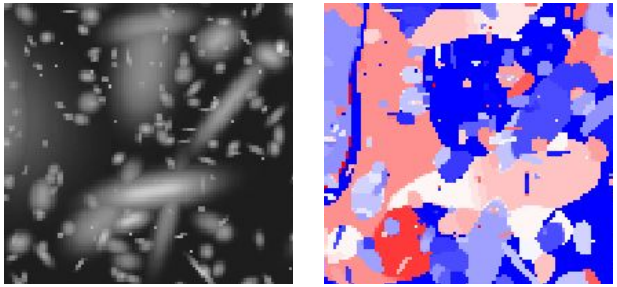






Future work

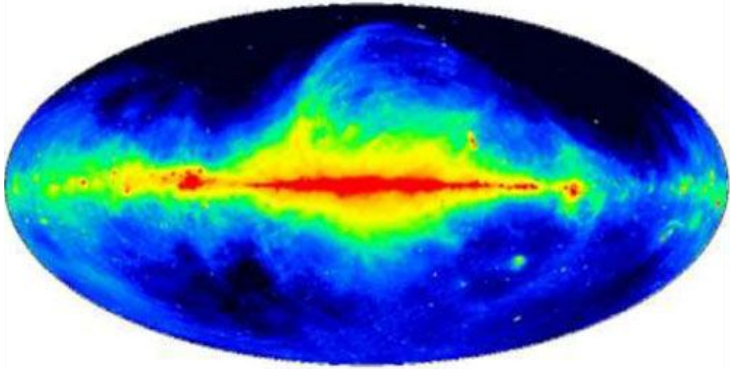
- Generalise to range of RMSFs
- Work in three dimensions
  - Removes independence assumption
- More complicated simulated observations
- Train on POSSUM early observations

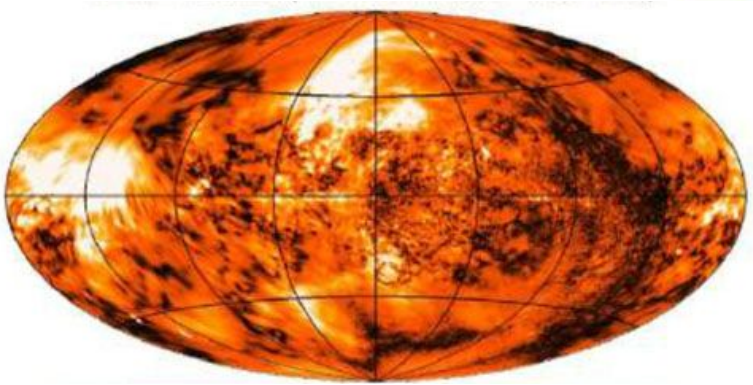


Simulated polarised sources and corresponding Faraday depths (Matthew Whiting).

Polarisation

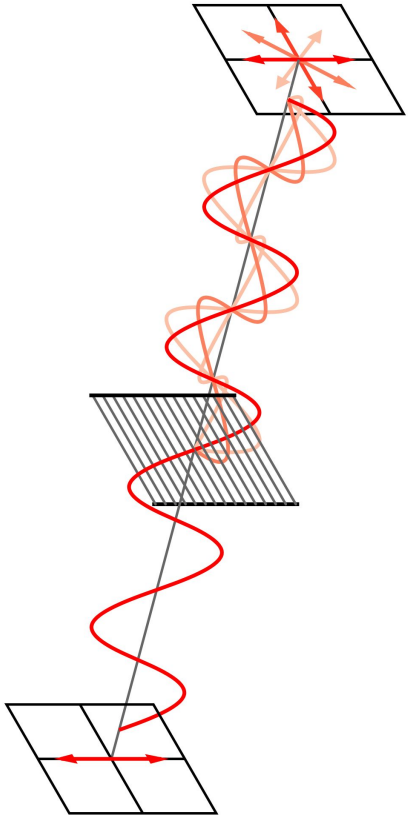
Total radio intensity.





Polarised intensity.

(Reich 1982; Wolleben et al. 2006; Testori et al. 2008).



Polarised waves. (http://physicsopenlab.org)