# A Fast Framework for Robust Decomposition of Quasar Spectra

Liam de Búrca

*Research Assistant, DARK, Niels Bohr Institute, University of Copenhagen*

*Supervisor: Prof. Marianne Vestergaard*

## Context and Motivation: Developing a Standardised Approach to AGN Modelling

Modern high-volume active galactic nuclei (AGN) surveys are integral to understanding the nature of AGN and their impact on host galaxies. Current surveys are projected to obtain on the order of a million spectra or more, allowing researchers to test hypotheses more rigorously. Due to the inherent complexity of these spectra and the known dependence of results on subjective methods [1], I am developing a high-performance modelling framework for analysing large volumes of spectra. This framework has four main branches containing data preparation routines, models, uncertainty estimation utilities, as well as other utilities that improve the ease-of-use and reliability of framework.

## Data Preparation

Prior to any modelling, raw data is prepared to streamline downstream processes. This includes:

- Redshift correction
- Logarithmic rebinning to a constant **velocity dispersion**
- Milky Way reddening correction
- **Narrow absorption line** (NAL) identification using signal smoothing techniques and a **structure-sensitive** normality test based on a **fast Fourier transform** (FFT)

## Modelling

The framework is built on **custom** ‘**Astropy**’ models:

- Power law continuum
- Adaptive emission line fitting using Gaussian functions and machine learning
- Numerically efficient template fitting of:
  - Iron pseudo continuum
  - Balmer continuum and series
  - Host galaxy emission

## Uncertainties

Following a successful fit, the user can choose to run uncertainty estimation routines, such as:

- Spectrum (w/o refitting) or parametric bootstrapping.
- Nested sampling using the **‘Dynesty’** module

The output is passed to a dedicated module that efficiently and robustly extracts model- and/or spectrum-based measurements using compiled ‘**Cython**’ code.

## Pipeline & Utilities

The framework includes quality-of-life features that make modelling large samples of spectra easier:

- I use the Python module ‘**Pydantic**’, a validation code used by e.g., ‘OpenAI’ and ‘Amazon’, to ensure that all input data and intermediate variables are strictly validated.
- I created various **I/O** routines to make sure that output, figures, and log files are correctly organised.

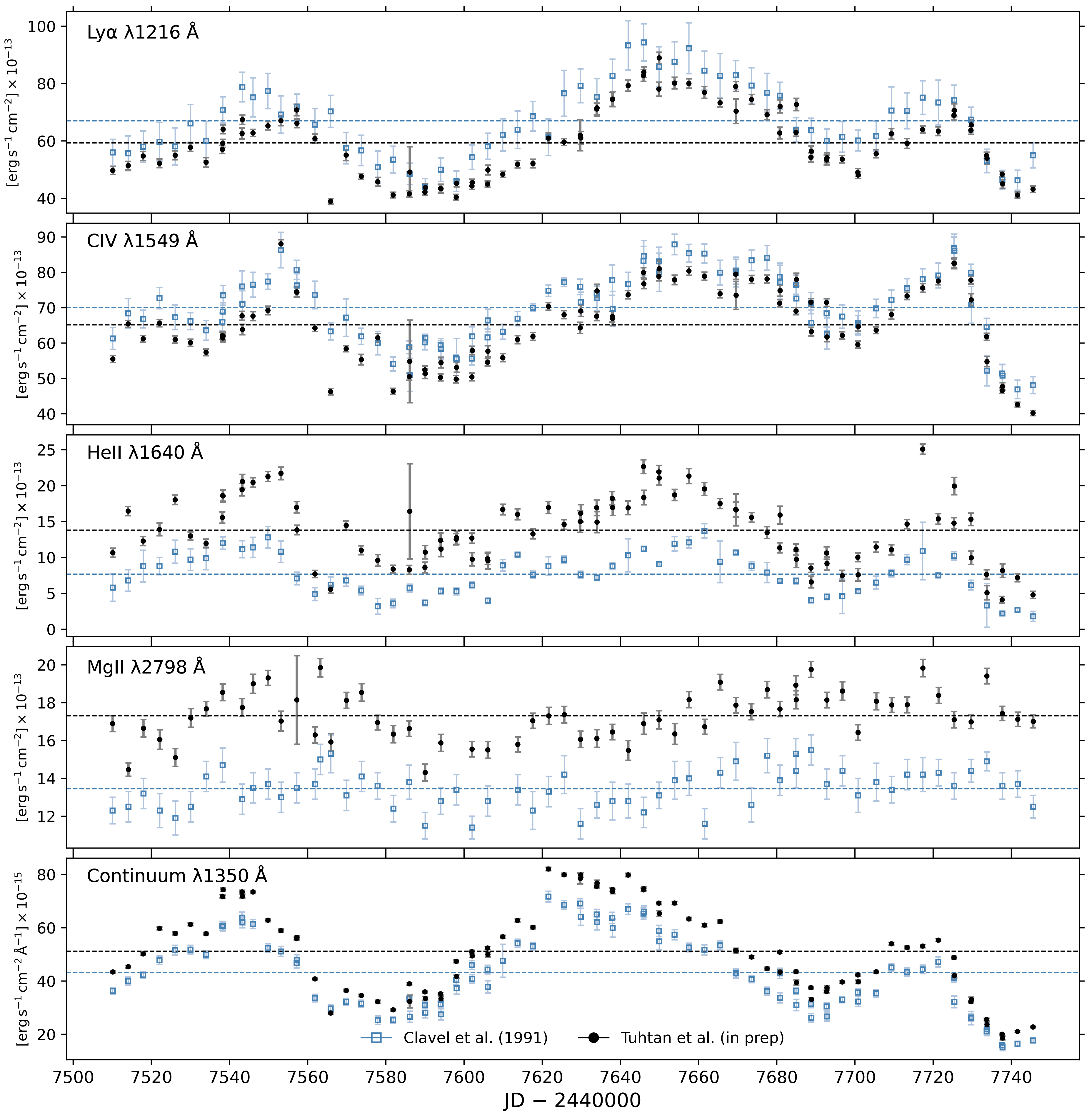

**Fig. 3:** Light curves from NGC 5548. Results gathered using my framework are shown in blue while results from [4] are shown in black. Credit: [3].

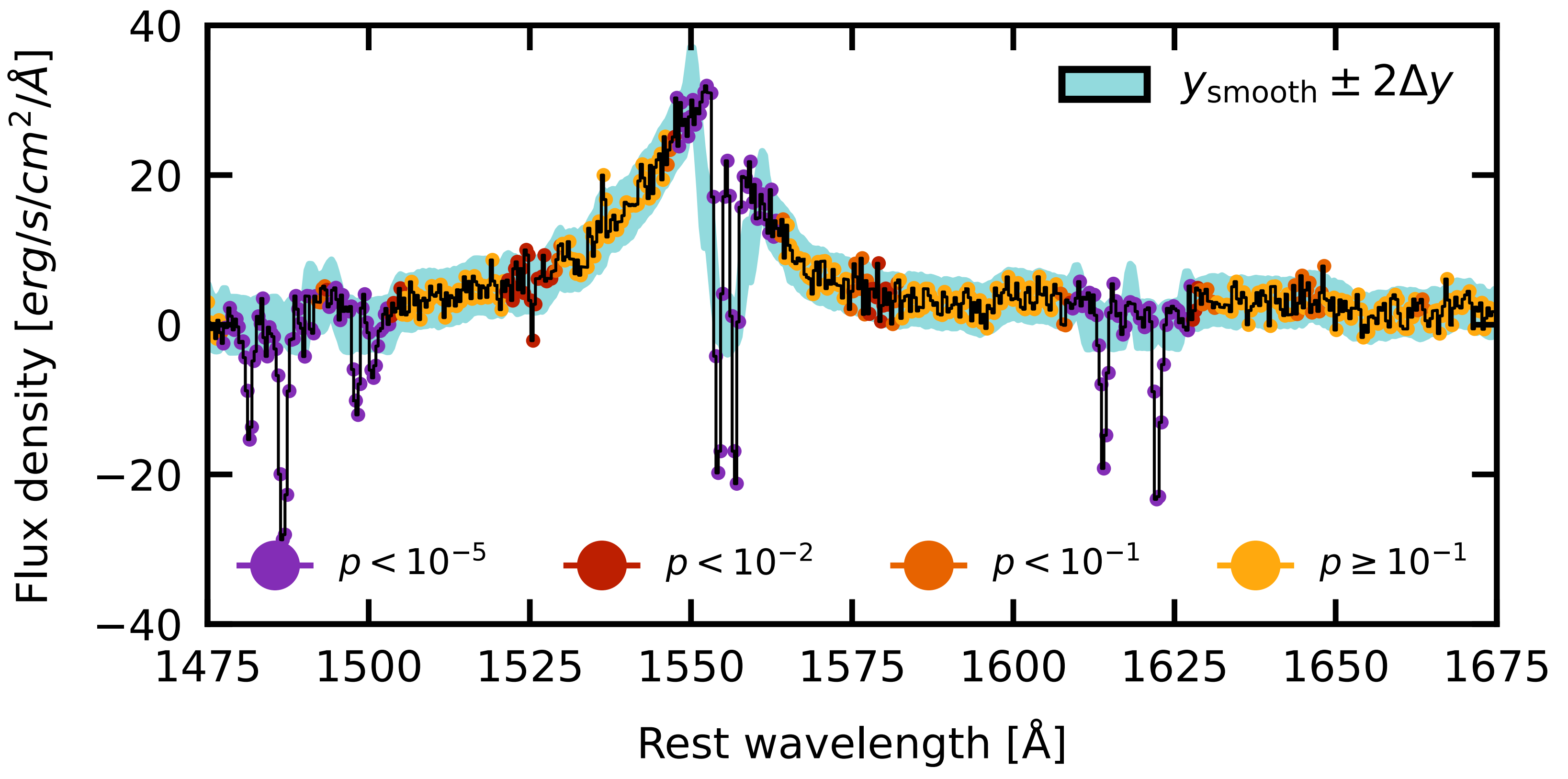

**Fig. 1**: Example of NAL identification algorithm using an **FFT-based** normality test on SDSS DR 7 data [2]. Purple points indicate highly anomalous pixels associated with systematic structure in the residuals, such as NALs.

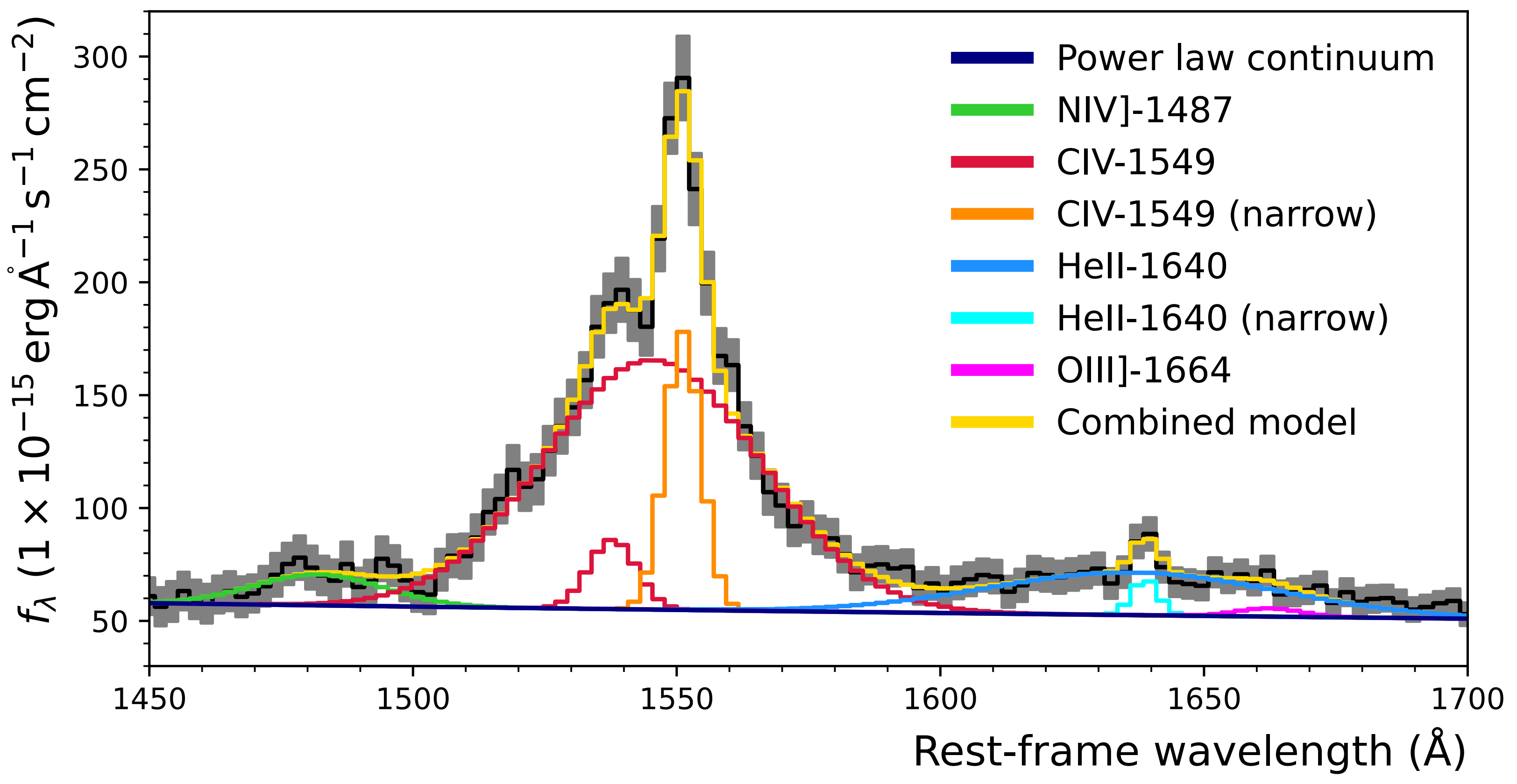

***Fig. 2:** Example of the adaptive Gaussian modelling technique on IUE data [4]. The narrow components of C IV λ1549 and He II λ1640 are shown in orange and cyan, respectively. The grey shaded area covers the $\pm 2\sigma$ interval.*

## Science Application: Light Curves

I have applied my pipeline to the IUE monitoring dataset [4] of AGN **NGC 5548** to measure the line and continuum fluxes (Fig. 3). These light-curves will be analysed to explore the **photoionisation physics** of the broad line emitting region [3]. I first modelled the mean stack of the 140 spectra to optimise the general model and then adjusted the solution to the spectrum taken at individual epochs.

This approach improves upon previous attempts [4] by:

- using more advanced emission line velocity profiles as well as modelling and removal of narrow line emission,
- modelling the blue wing of **Lyα**, partially covered by **geocoronal emission**, and
- applying an iron pseudo-continuum template surrounding the **Mg λ2798** doublet.

## Future Science Application: 4G-PAQS

The **4MOST-Gaia Purely Astrometric Quasar Survey** (4G-PAQS) will provide **100,000** quasar observations selected purely using Gaia astrometry to improve our understanding of physics and selection effects associated with intervening absorbers along the line of sight.

As a member of the 4MOST S6 team I am developing a **first-look pipeline** based on this framework to distinguish between different types of sources and identify spectra worth more detailed modelling, e.g. those of broad absorption line quasars (BALs).

*References:*
*[1] Vestergaard et al. 2011, DOI 10.22323/1.126.0038;*
*[2] Abazajian, et al. 2009, DOI 10.1088/0067-0049/182/2/543;*
*[3] Tuhtan et al. (in prep);*
*[4] Clavel et al. 1991, DOI 10.1086/169540*