Northwestern
CIERA
CENTER FOR INTERDISCIPLINARY EXPLORATION AND RESEARCH IN ASTROPHYSICS

LS4

SkAI
Funded by the U.S. National Science Foundation and the Simons Foundation

Galaxies as Friends and Foes

Morphological Classification and Spectral Decontamination of Galaxies for Time-Domain Astronomy

Chang Liu, Adam A. Miller CIERA/Northwestern

2025.6 Berkeley

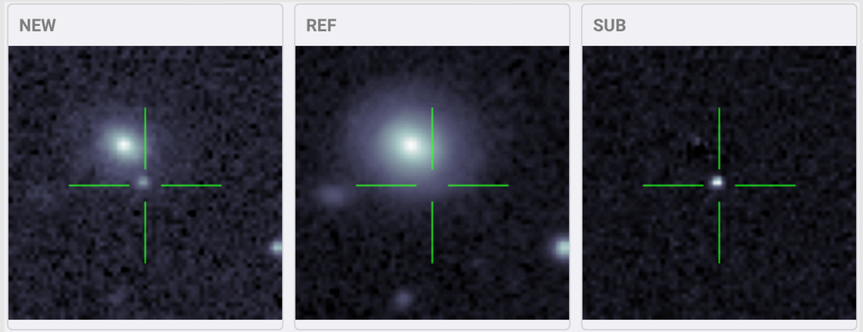
Part 1

A Morphological “Star—Galaxy” Classifier for the DESI Legacy Surveys

Star-Galaxy Classifier in the Alert Stream

A dichroic for Galactic/extragalactic transients

Data flowing in



→ **Pass**

(i) Is there a cataloged object associated with the transient?

(ii) Is it a star or galaxy?

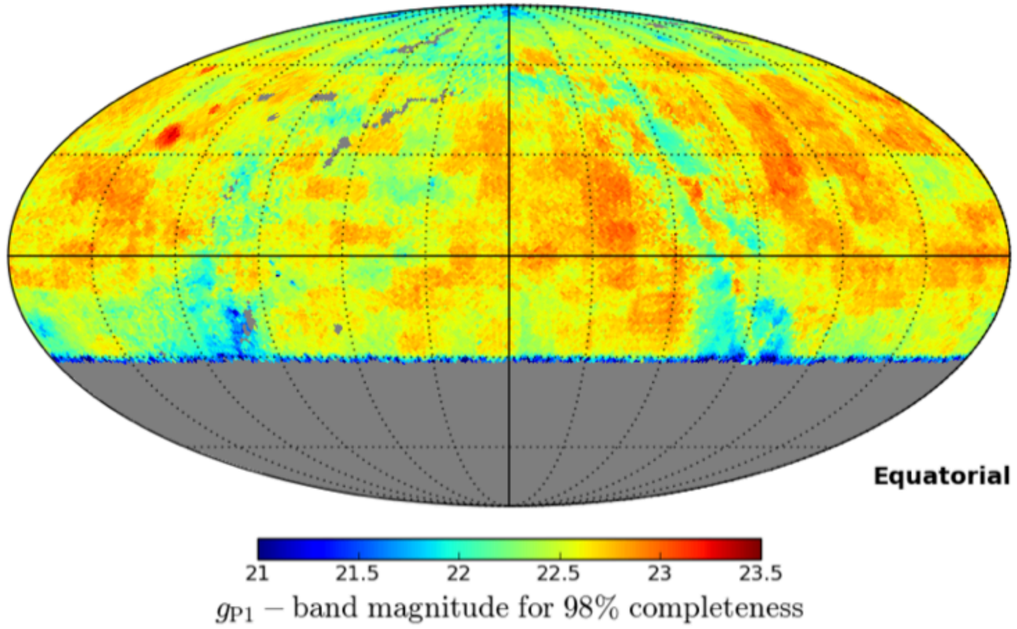
→ **Fail** (stars)

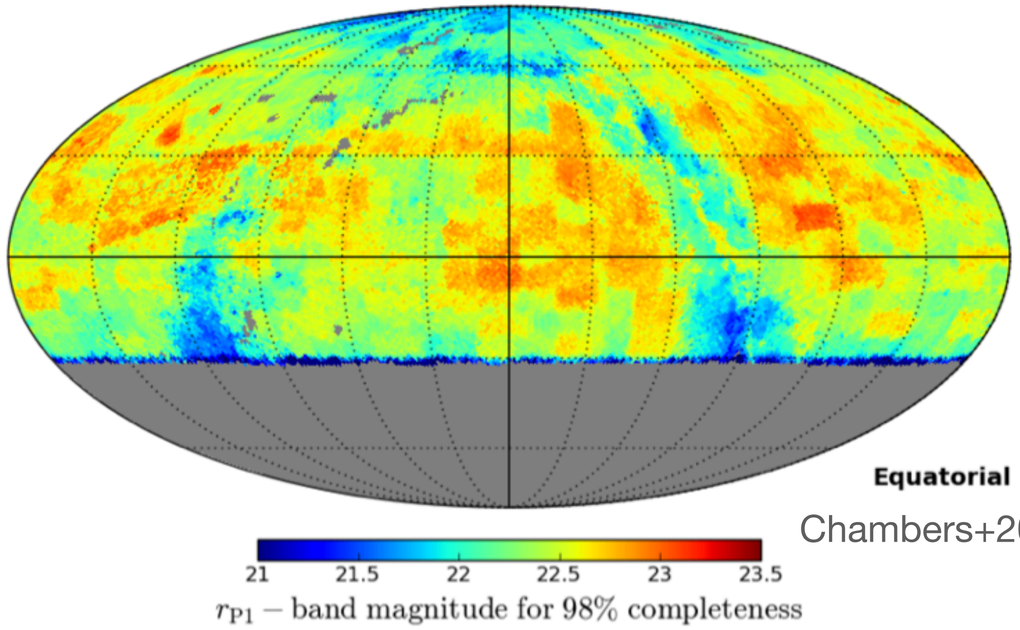
To the next filter

SGSCORE

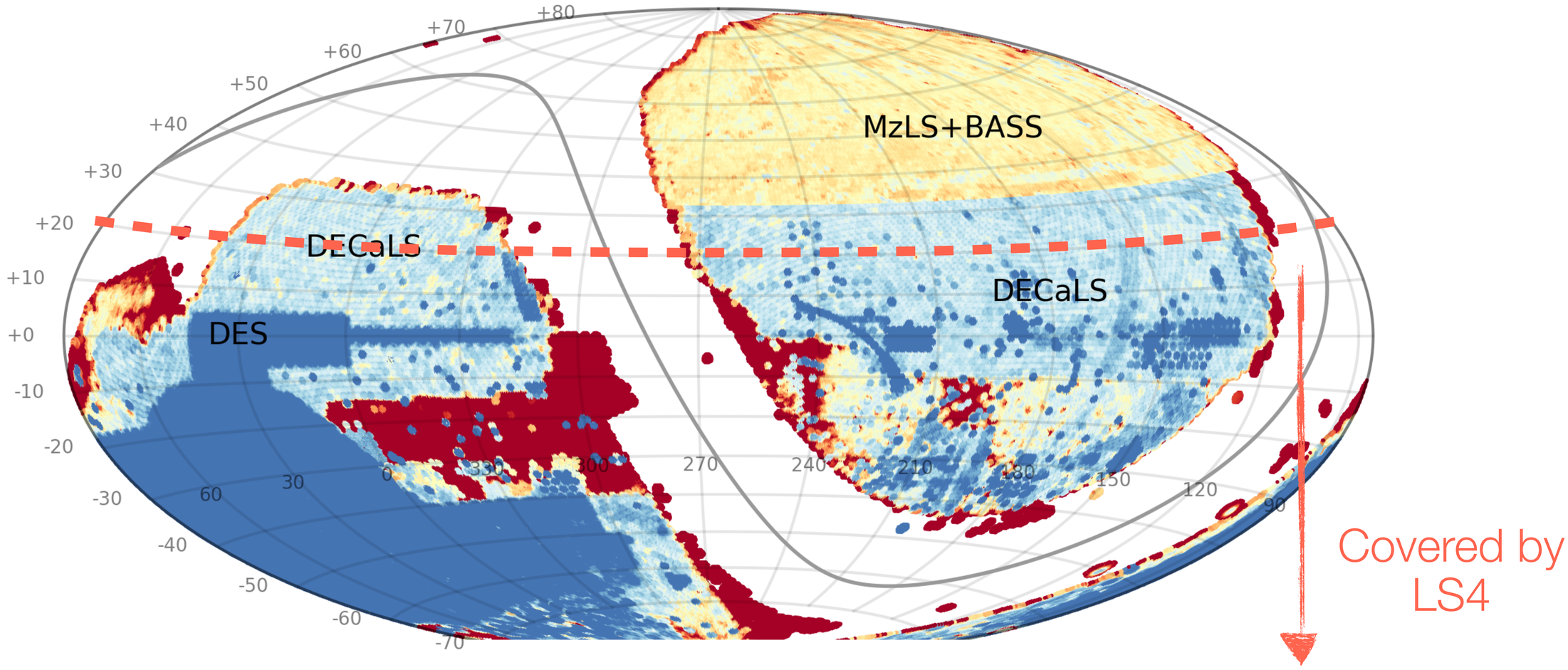
Tachibana & Miller 2018

- Implemented in the **Zwicky Transient Facility (ZTF)** real-time pipeline
- Based on **Pan-STARRS** 3π survey
  - Morphological classifier — **photometry**
  - Random forest — **lightweight & efficient**
  - Does not cover the south — **dec > -30°**
  - Limited by depth — **≲ 23 mag**

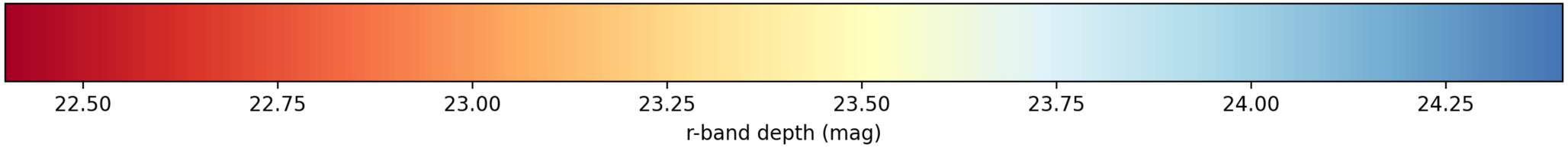




Chambers+2016

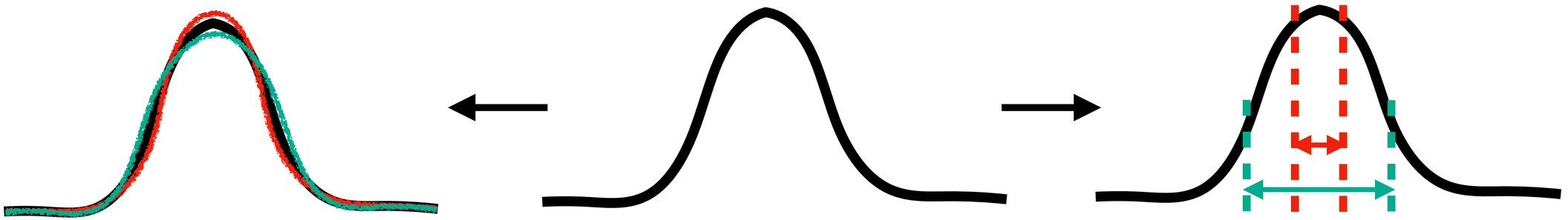


Legacy Surveys (LS) DR10 — 3 billion sources



Star-Galaxy Classifier for LS4

Morphology-based features



Point source — PSF
Galaxy profile models

Observed profile

Aperture 1
Aperture 2, …

dmlc XGBoost

Chen & Guestrin 2016

Feature group 1
How well each model fits the profile

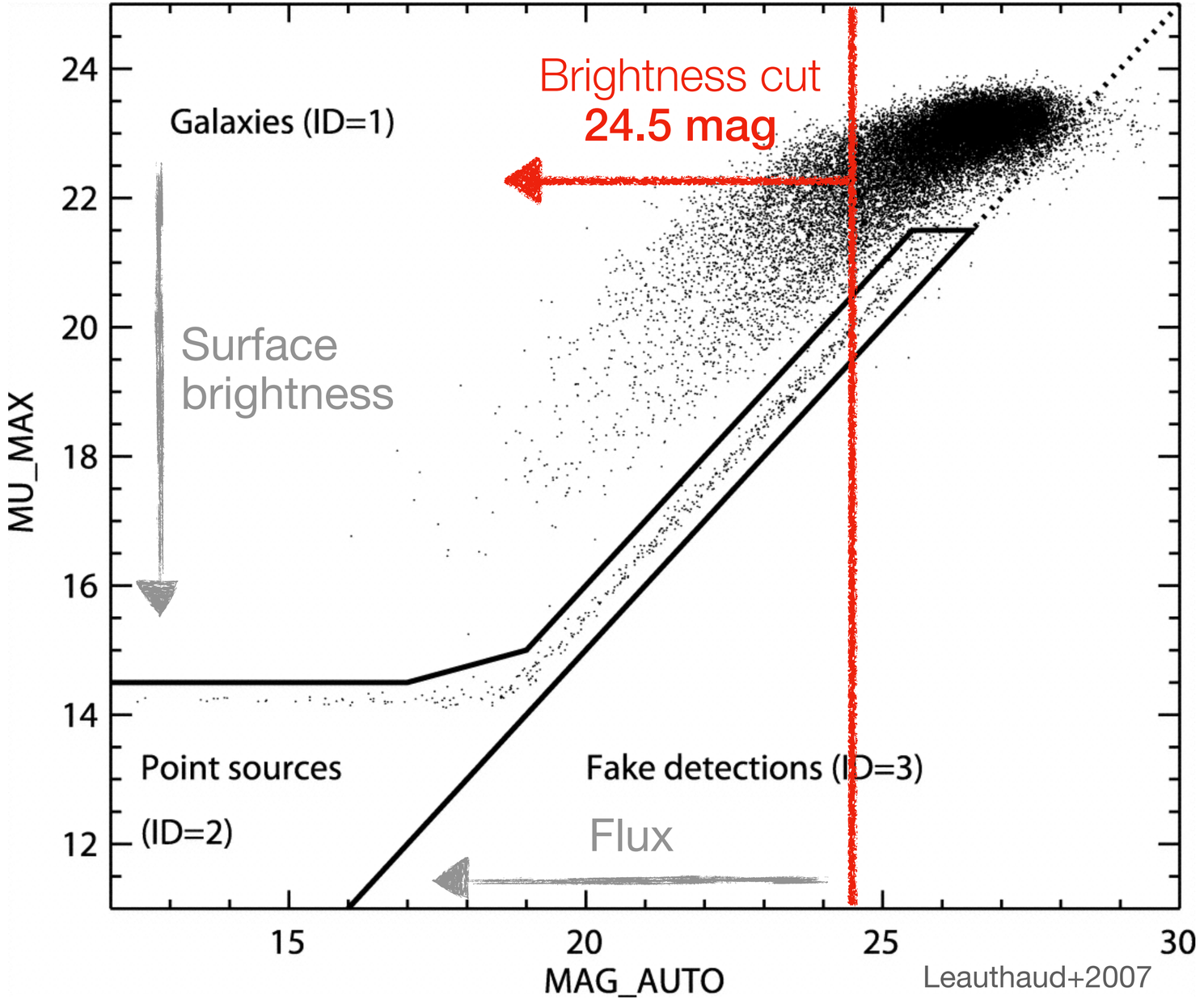
Feature group 2
How concentrated the flux distribution is

Resolved/unresolved score

LS X HST

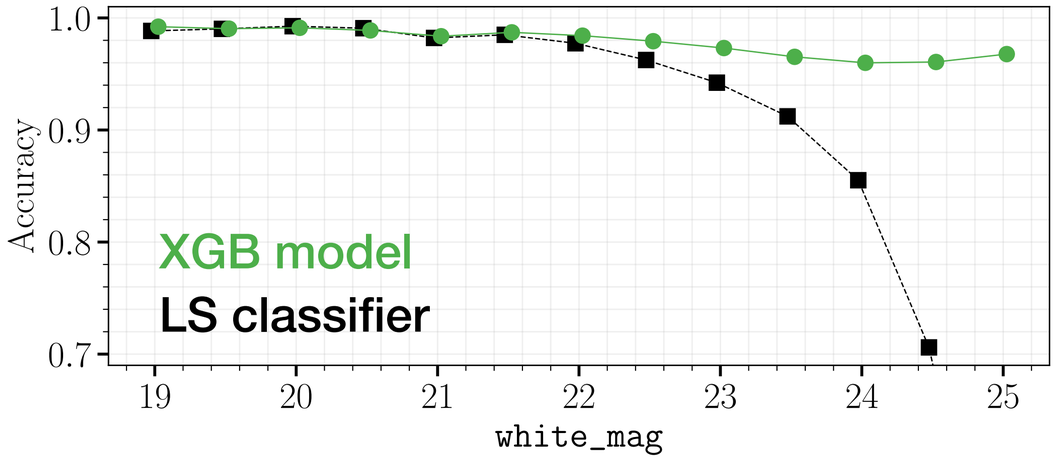
Training set

- COSMOS field
  - Deep images from both LS and Hubble
  - Confident labeling down to **~24 mag**
  - **~240,000** objects in the training set



Performance

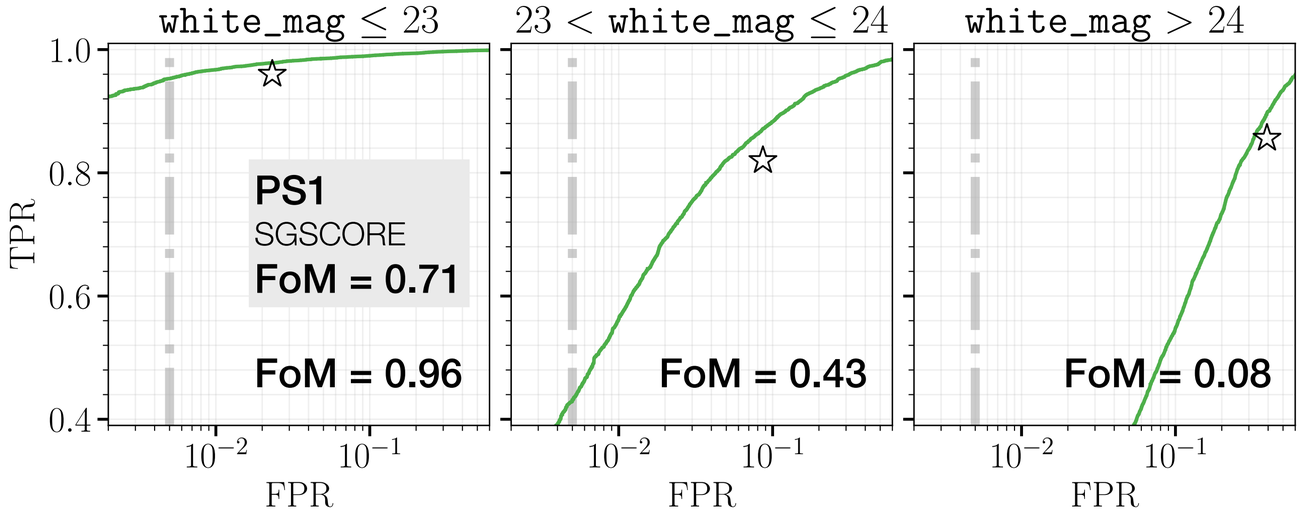
Cross validation



High overall accuracy

Accuracy = 0.97

Fraction of correctly classified stars + galaxies



Sensitive to faint galaxies

Figure of merit (FoM) = 0.71

Fraction of recalled stars when misclassifying 0.5% of galaxies

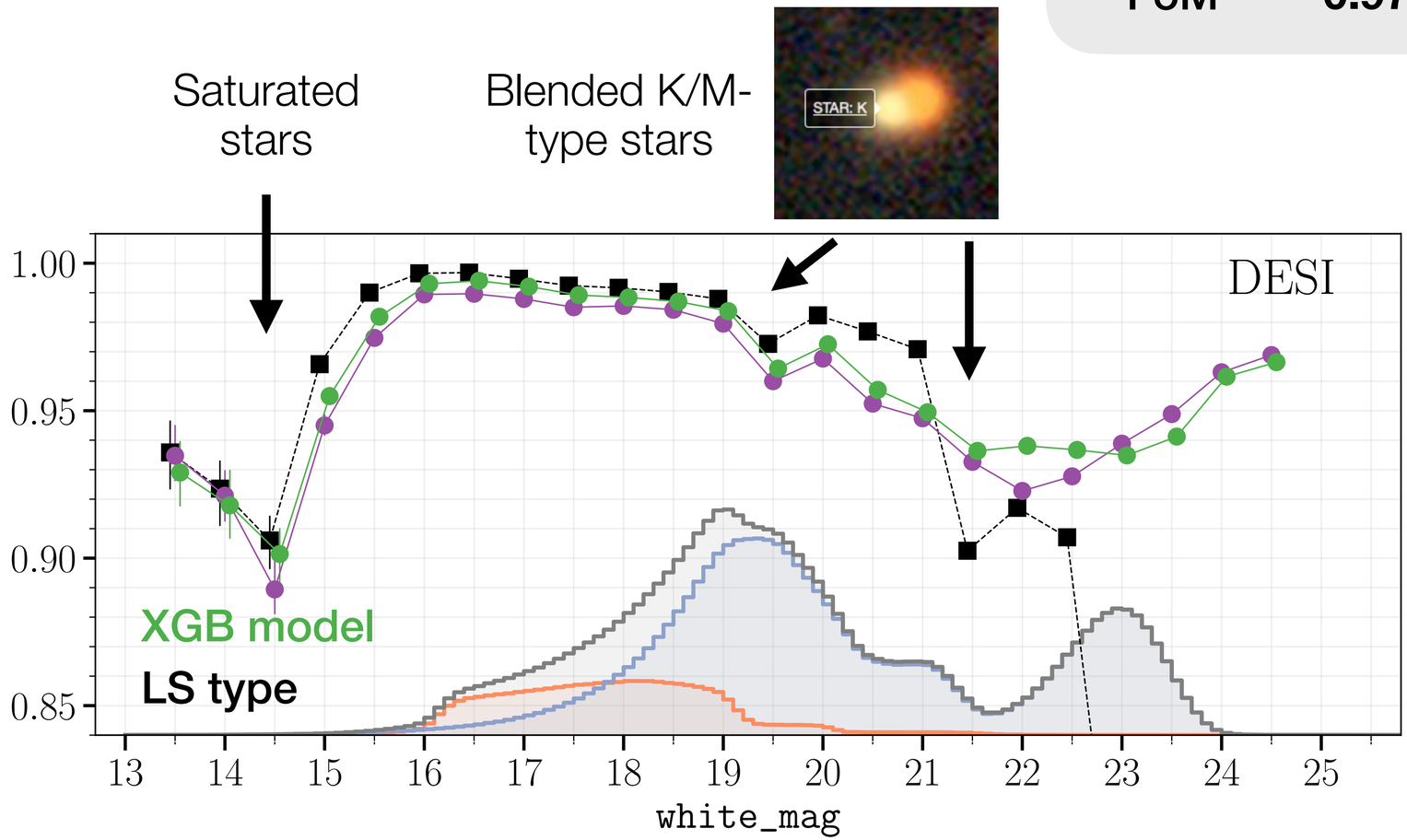
Performance

Test on spectroscopically classified samples

DESI DR1 (15,354,071)	XGB	LS
Accuracy	**0.9919**	0.9451
FoM	**0.9778**	…

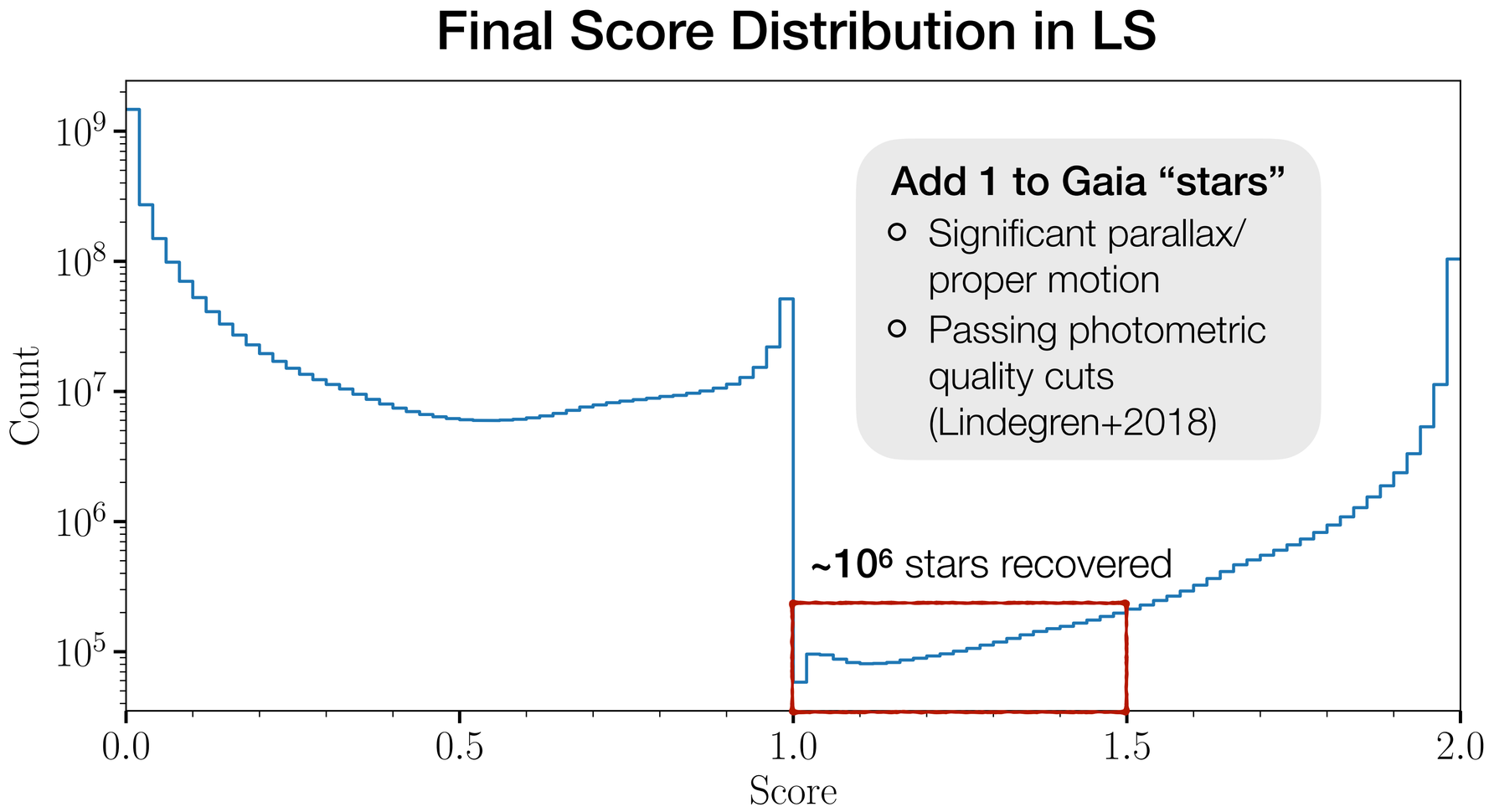
Normalized Accuracy

The star/galaxy ratio in each bin is tweaked to match the COSMOS field



Systematics (Saturated Stars)

Leveraging Gaia DR3



Summary

LS Point Source Catalog

- A **morphological** classifier based on LS photometry trained with XGBoost
- Outperforms the **LS classifier** (extra features + flexible ML modeling) & **PS1 SGSCORE** (deeper images)
- Near-perfect classification for targets ≲ 23 mag
- Optimized for keeping galaxies on the faint end

Liu et al., 2025 (arXiv:2505.17174)

https://ls-xgboost.lbl.gov/ (**Thanks to Rob Knop!**)

Part 2

Precise Galaxy Light Subtraction in Transient Longslit Spectroscopy

Image Subtraction for Spectroscopy?

Removing host contamination

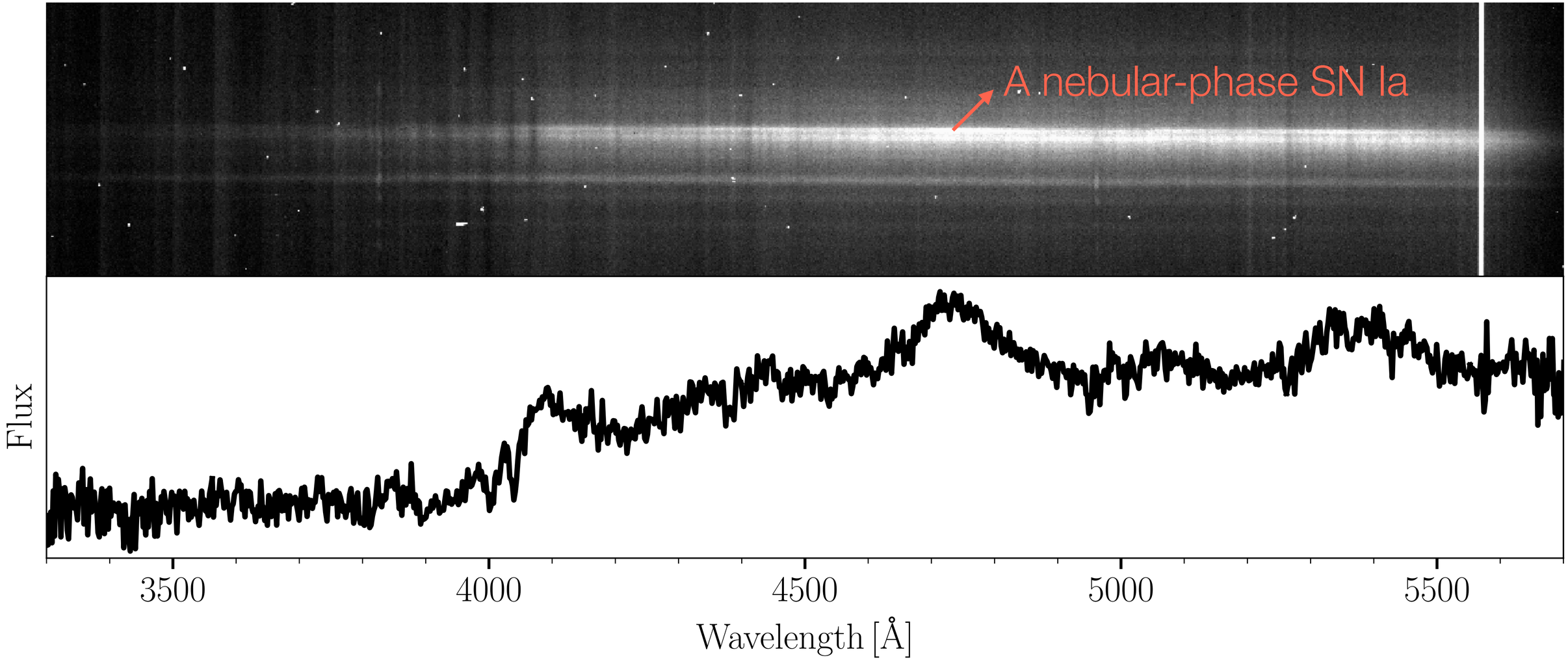
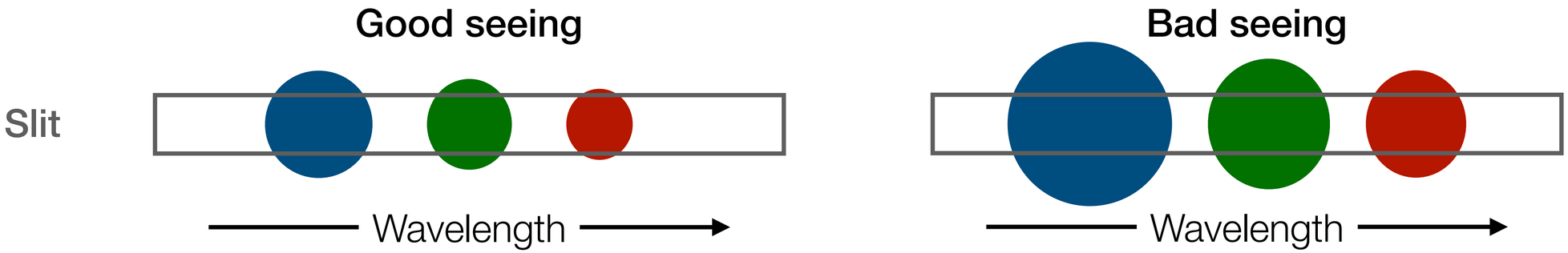


Image Subtraction for Spectroscopy?

Difficulties



- **Observing condition:** seeing, sky background…
- **Instrumental configuration:** slit width, position angle…
- **Reference spectrum not accessible:** time for large mirrors is always expensive!

README MIT license

HostSub_GP

Liu et al., 2025b, in prep.

Modeling the 2d spectrum of host galaxies with Gaussian process (GP) for better background subtraction in supernova spectroscopy.

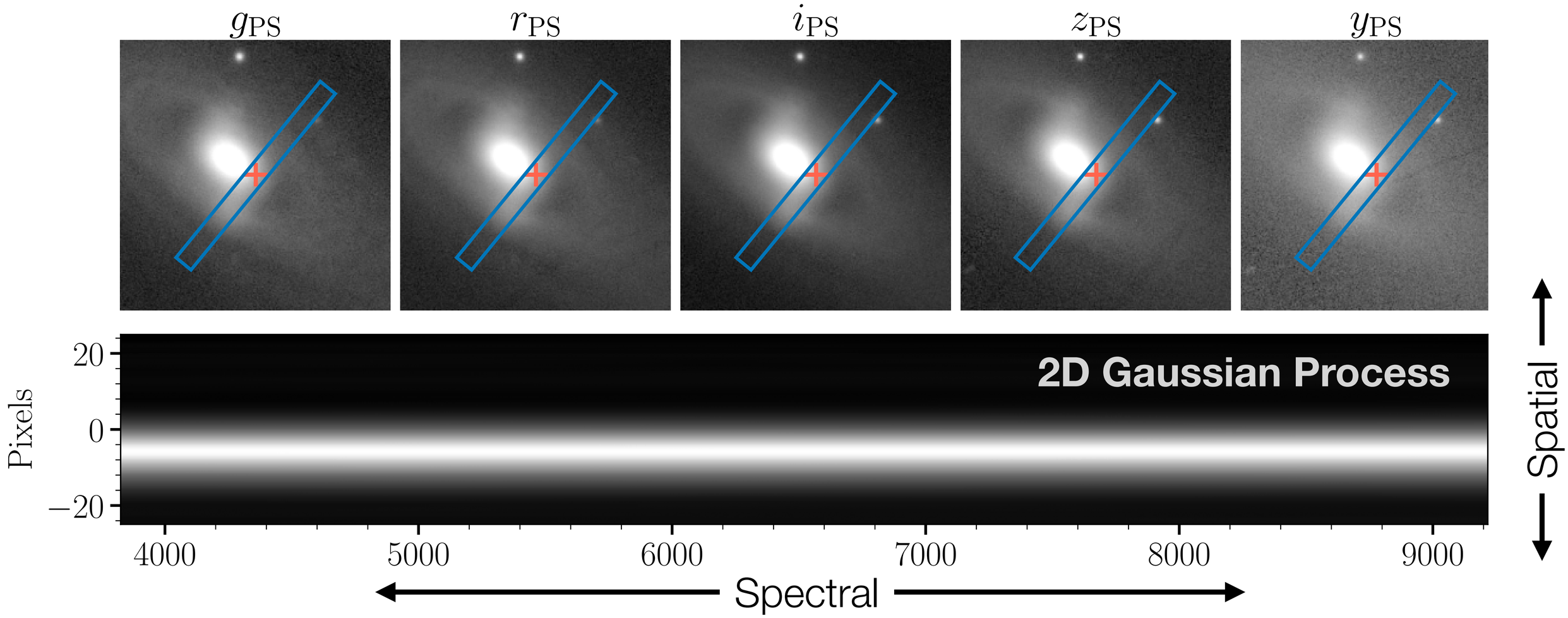
- Fully **Python** + open source (soon!)
- Leveraging archival imaging as **priors**
- Accelerated by JAX: **Just-in-time** compilation + **automatic differentiation**

Robust galaxy background removal in a few minutes on your laptop!

HostSub_GP Pipeline

Building the prior - normalized 2D profiles

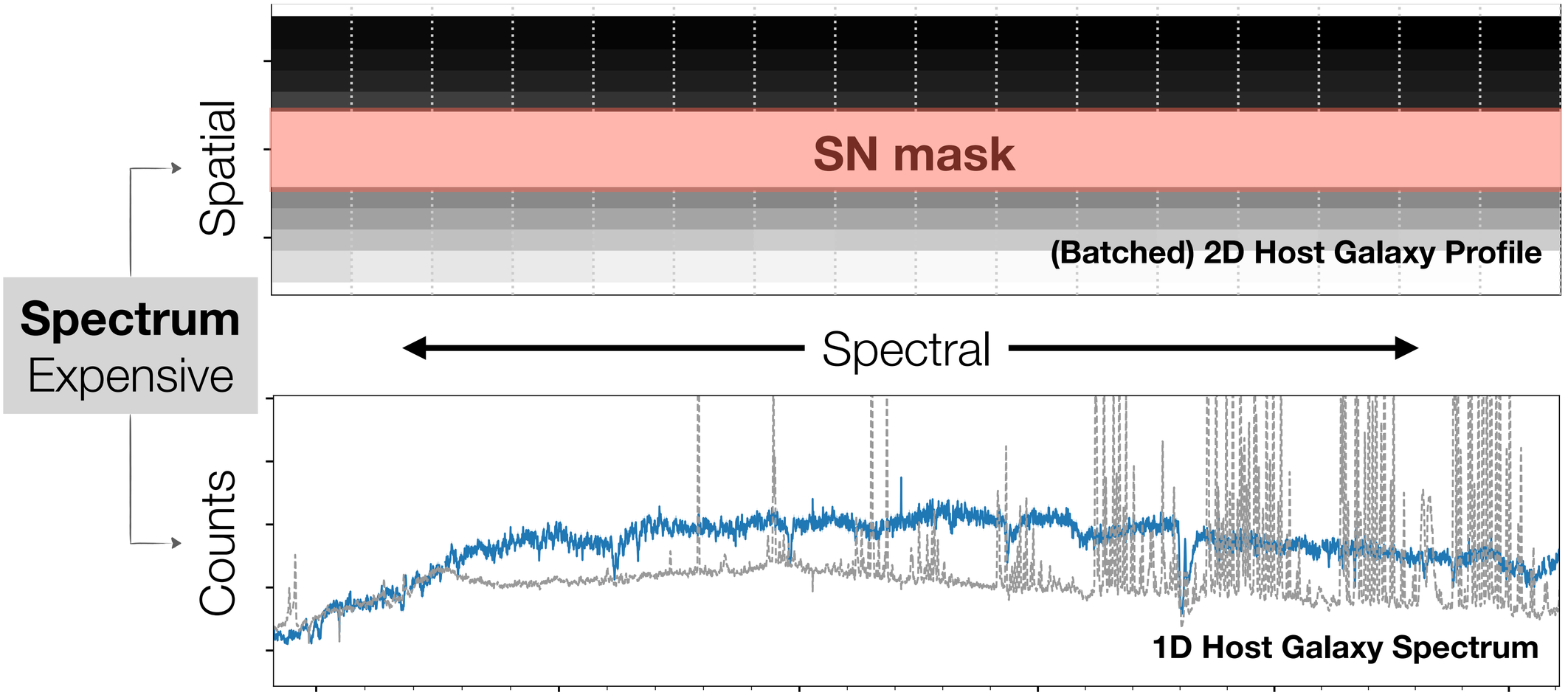
Liu et al., 2025b, in prep.



HostSub_GP Pipeline

Preprocess the 2D spectrum

Liu et al., 2025b, in prep.



2D GP

Model how the 2D spectrum deviates from the prior

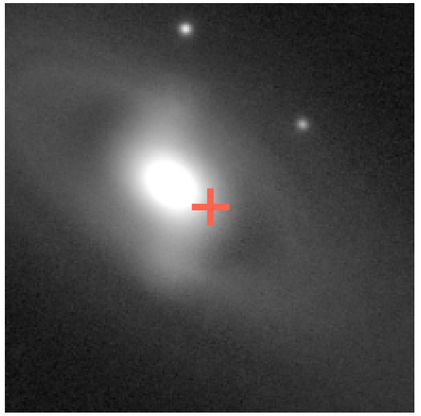
1D GP

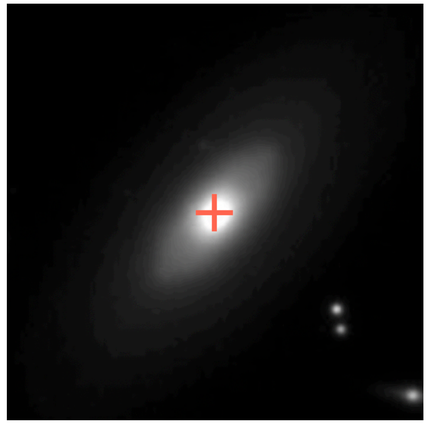
Model the 1D Spectrum

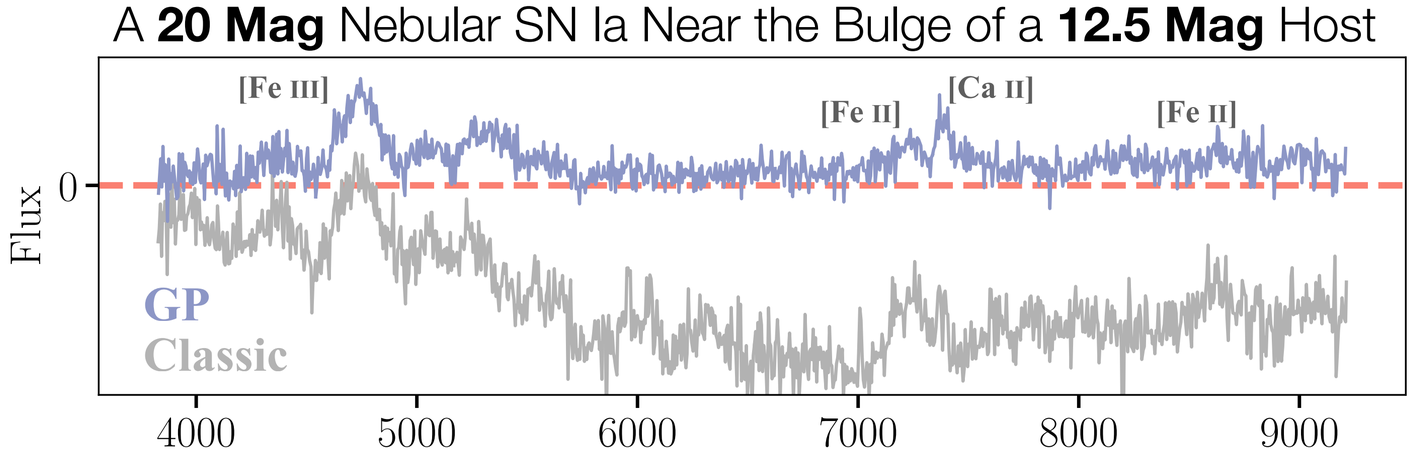
HostSub_GP Pipeline

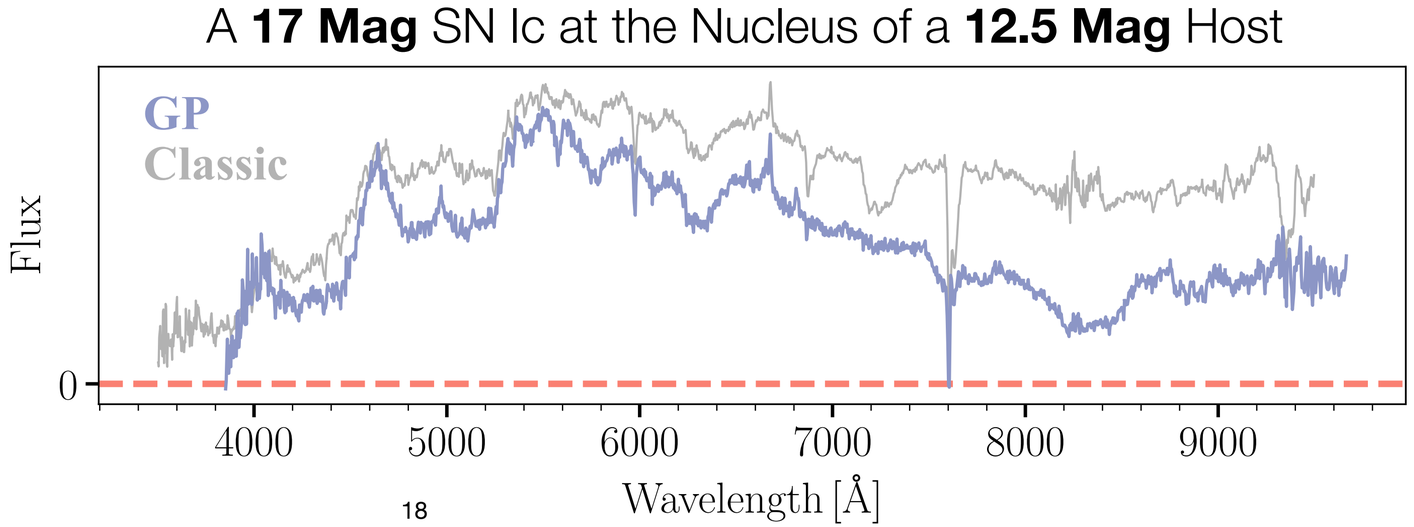
Useful Cases

Liu et al., 2025b, in prep.









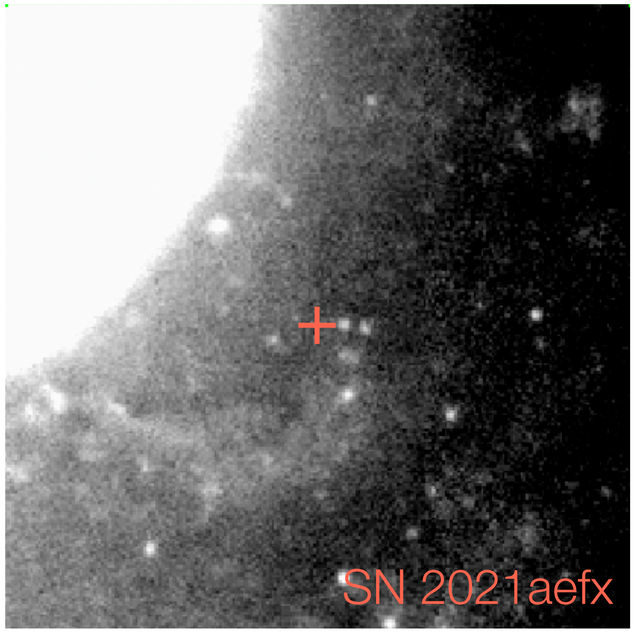
2025-04-30 20:19:16 Type: Announcement-General

Detection of a Light Echo from SN 2021aefx

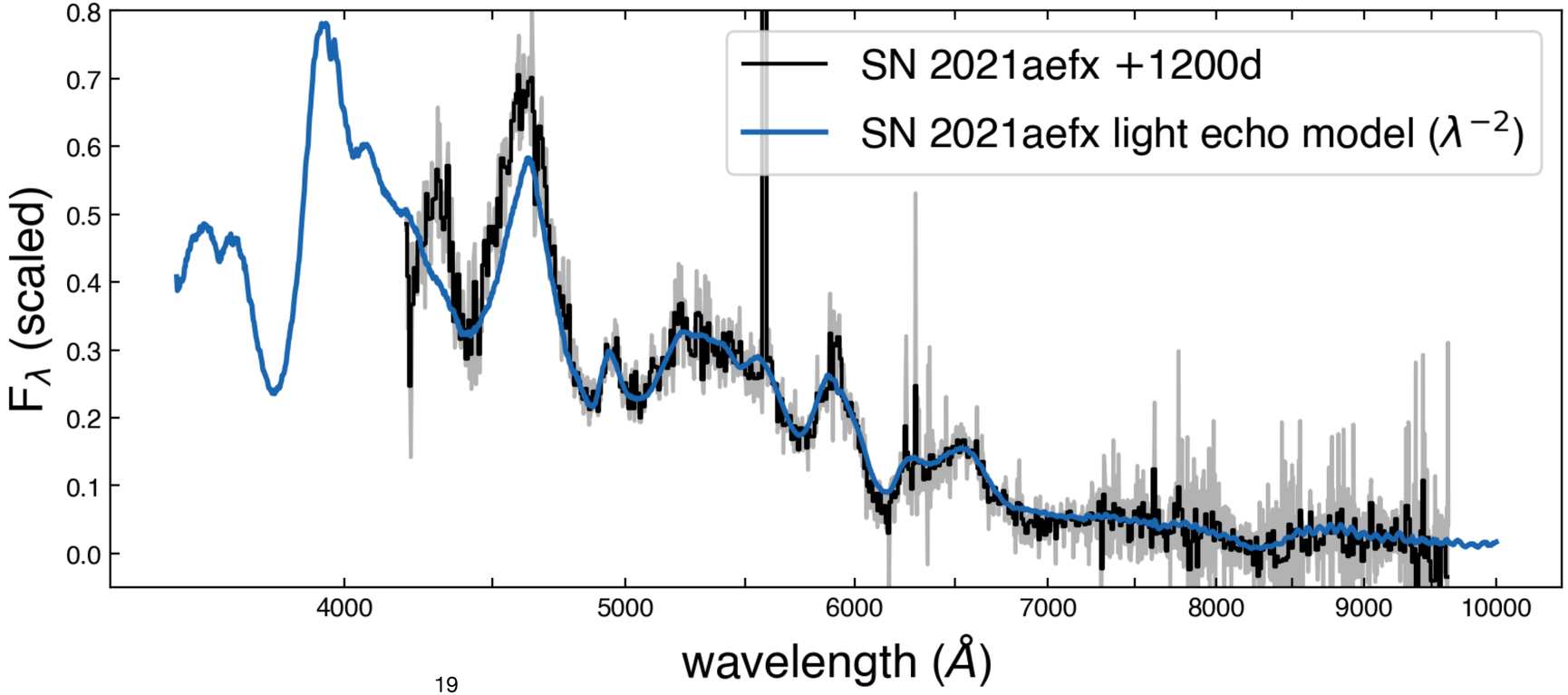
Authors: Lindsey A. Kwok (CIERA/NU), Steve Schulze (CIERA/NU), Chang Liu (CIERA/NU), Matthew R. Siebert (STScI), Joel Johansson (Stockholm), Huei Sears (Rutgers), Or Graur (University of Portsmouth), Griffin Hosseinzadeh (UCSD), Saurabh W. Jha (Rutgers), Adam A. Miller (CIERA/NU), David J. Sand (U Arizona), Stéphane Blondin (CNRS, LAM / ESO), Andreas Flörs (GSI), James M. DerKacy (STScI), Chris Ashall (U Hawaii), Peter Hoeflich (FSU)

Keywords: Supernova , Optical , Spectroscopy , Transient

Abstract: We report the spectroscopic detection of a light echo from Type Ia Supernova 2021aefx at ~1200 days post maximum light using the Very Large Telescope.

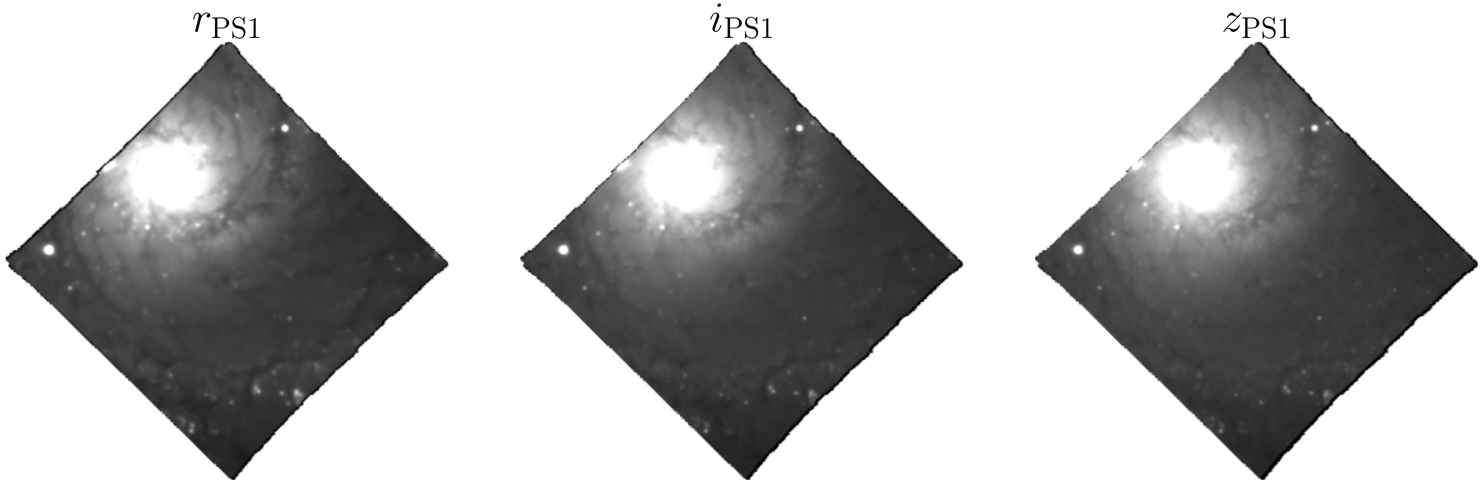


A **23.3 Mag** SN Ia (Echo!) With Complicated Galaxy Background

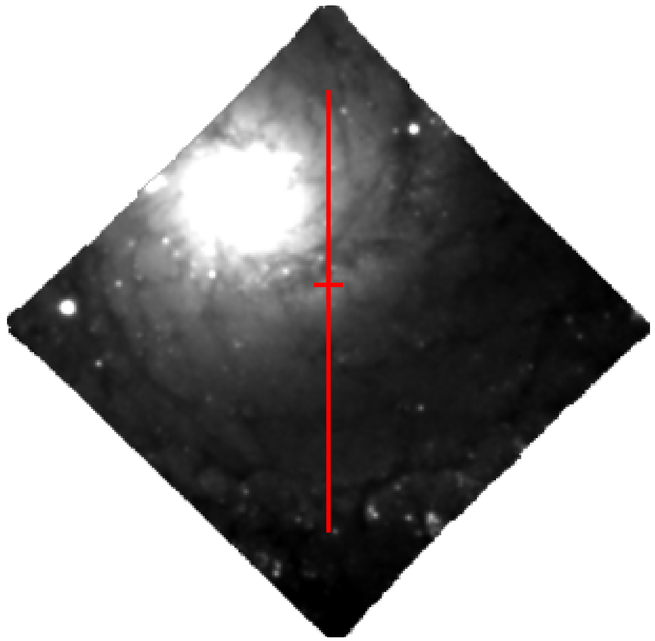


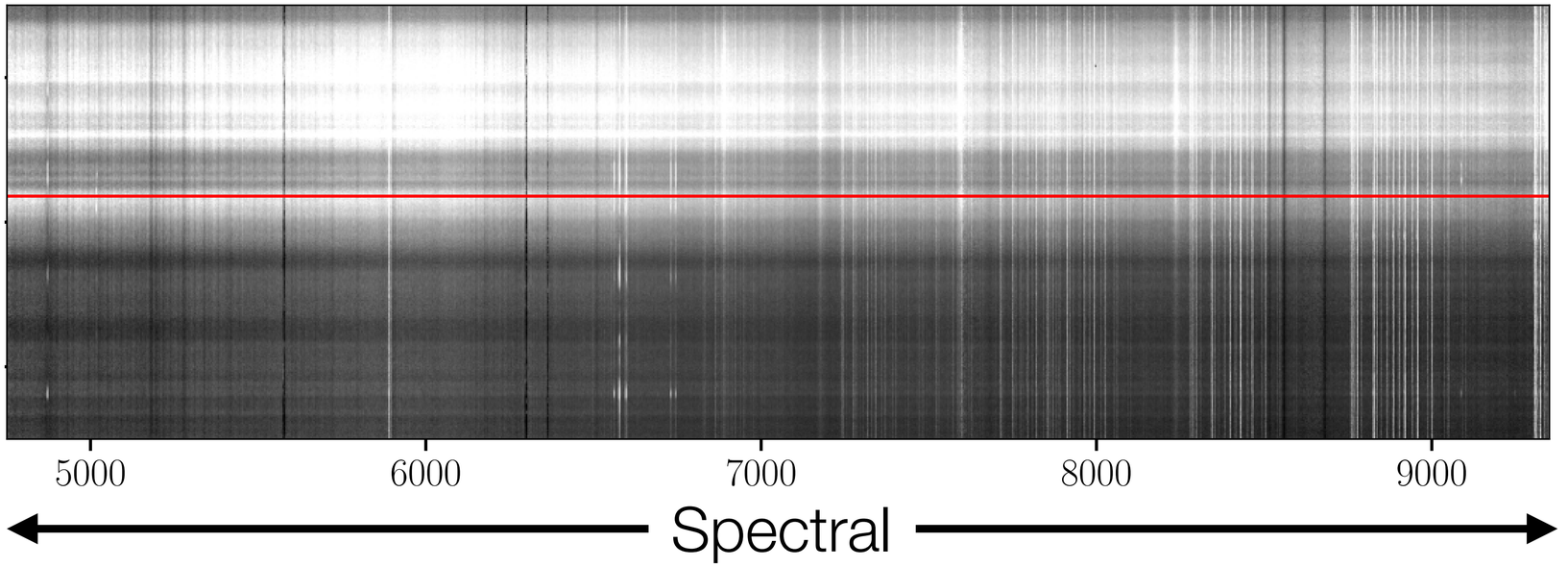
Test on a Synthetic Dataset

MUSE data cube



NGC 0628



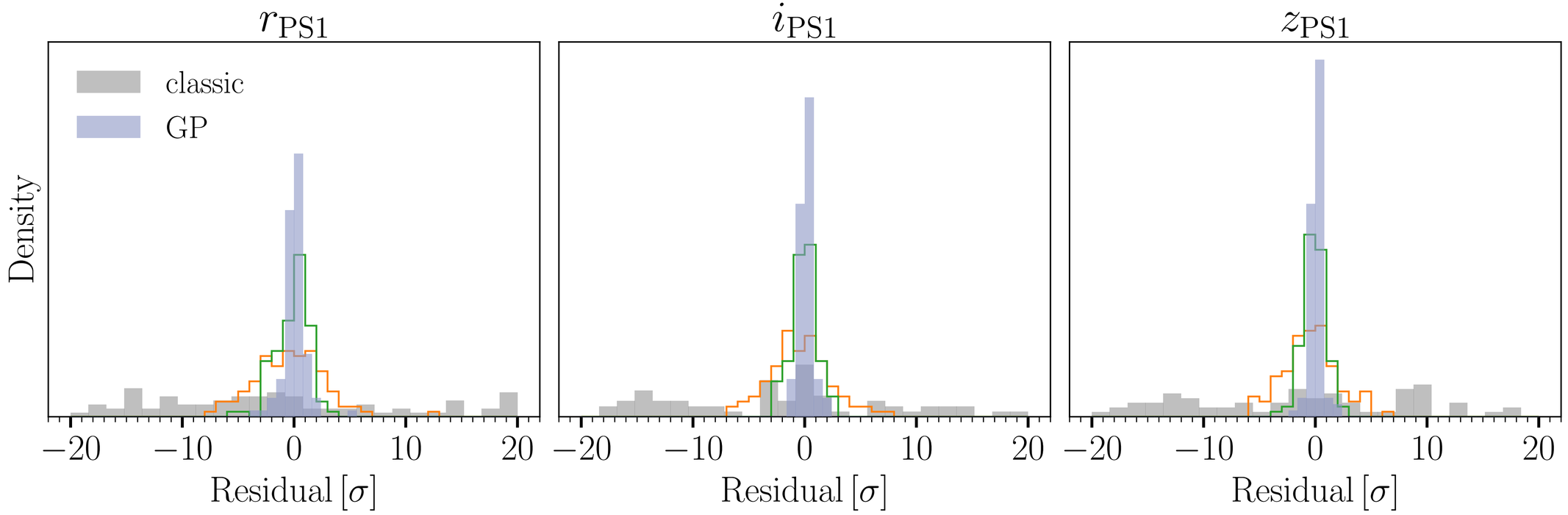


Test on a Synthetic Dataset

MUSE data cube: 100 different slit positions

Worsen the seeing in the synthetic image (0.7”—1.1”)

Seeing matched



Caveats

HostSub_GP can sometimes fail…

- Only works for host galaxies that are
  - **Well resolved:** containing sufficient pixels to be modeled
  - **Sufficiently bright:** for the separation of 1D flux and 2D profile
- Nebular lines cannot be robustly removed (flux profiles are usually very different from the prior)
- Adopting archival images from a mix of surveys is allowed but not recommended (inconsistent seeings)

Summary
HostSub_GP

- A toolkit to improve host subtraction in **SN longslit spectroscopy**
    - Not limited to SN - try your favorite **transient**!
- Leveraging archival photometric observations
- Modeling the galaxy background with 1D/2D Gaussian processes
- Outperforming the classic aperture spectroscopy on a synthetic dataset of a spiral galaxy with rich structures!