High-Resolution X-ray spectroscopy: A Chandra Workshop
August 1-3, 2023, Cambridge MA

CHALLENGES IN ANALYZING HIGH-RES SPECTRA OF WEAK SOURCES

VINAY KASHYAP (CfA/CHASC)

Sara Algeri and Xiangyu Zhang (Minnesota)
David van Dyk and Antoine Meyer (Imperial)
Margarita Karovska (CfA)

MINEFIELDS AHEAD

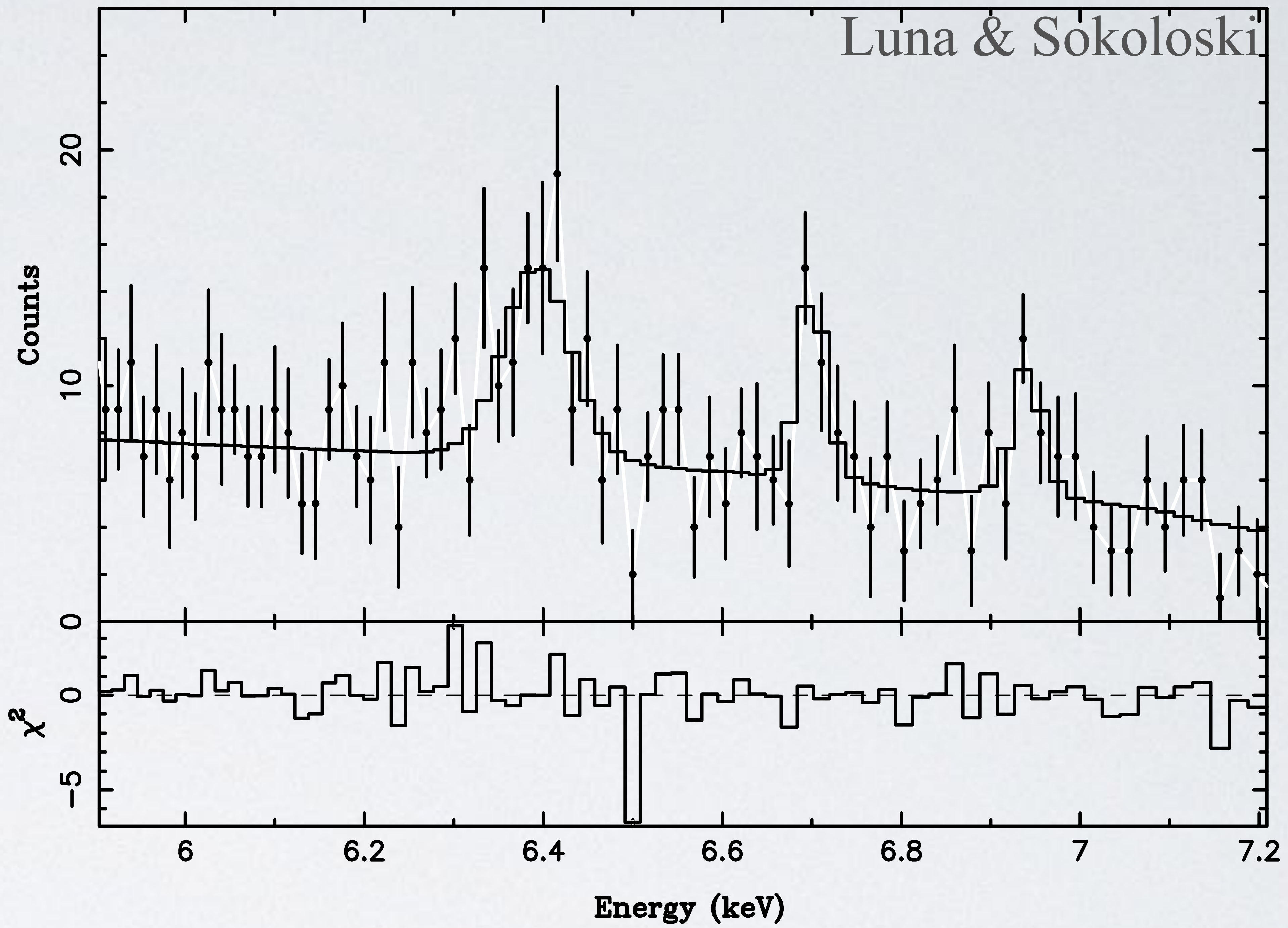
High-resolution spectra come with unintuitive challenges, caused by **sparsity**, **detectability**, and **background**

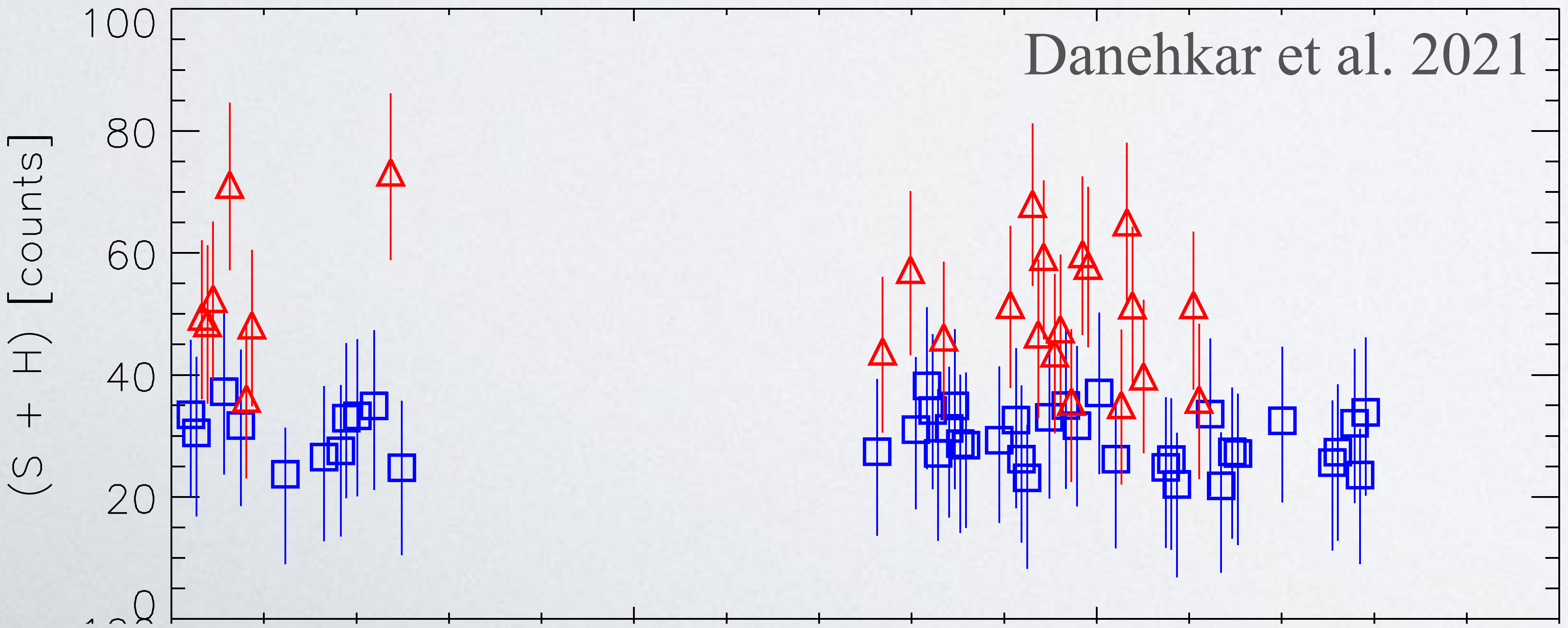
I. (RT Cru) Weak lines in high background

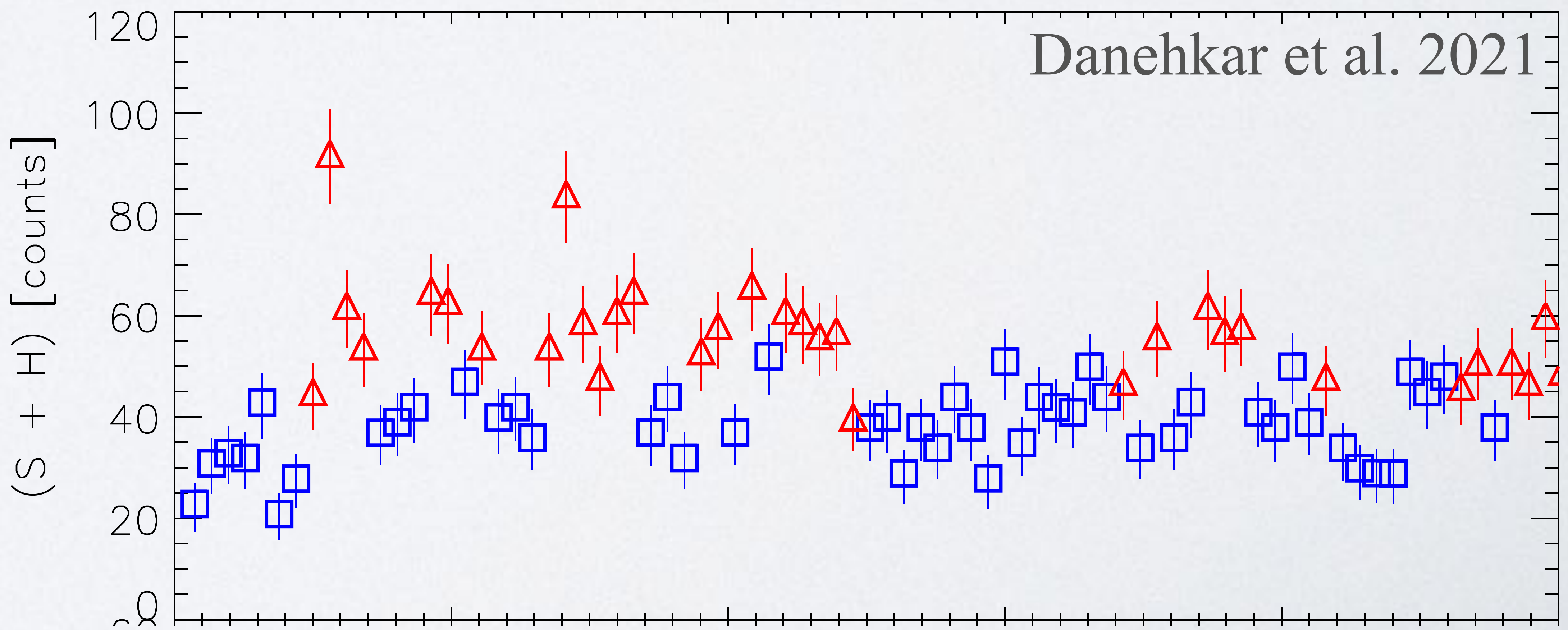
II. (UV Cet AB) Disentangling overlapping lines from contaminating companions

I. WEAK LINES IN HIGH BACKGROUND

- RT Cru is a symbiotic system at 2.5 kpc, with a high mass WD (1.3 $M_\odot$) accreting from a M5III giant
- Exhibits aperiodic flickering, with heavily absorbed hard power-law component, strong lines from Fe XXV, Fe XXVI, and FeKα, and possibly an absorbed soft thermal component

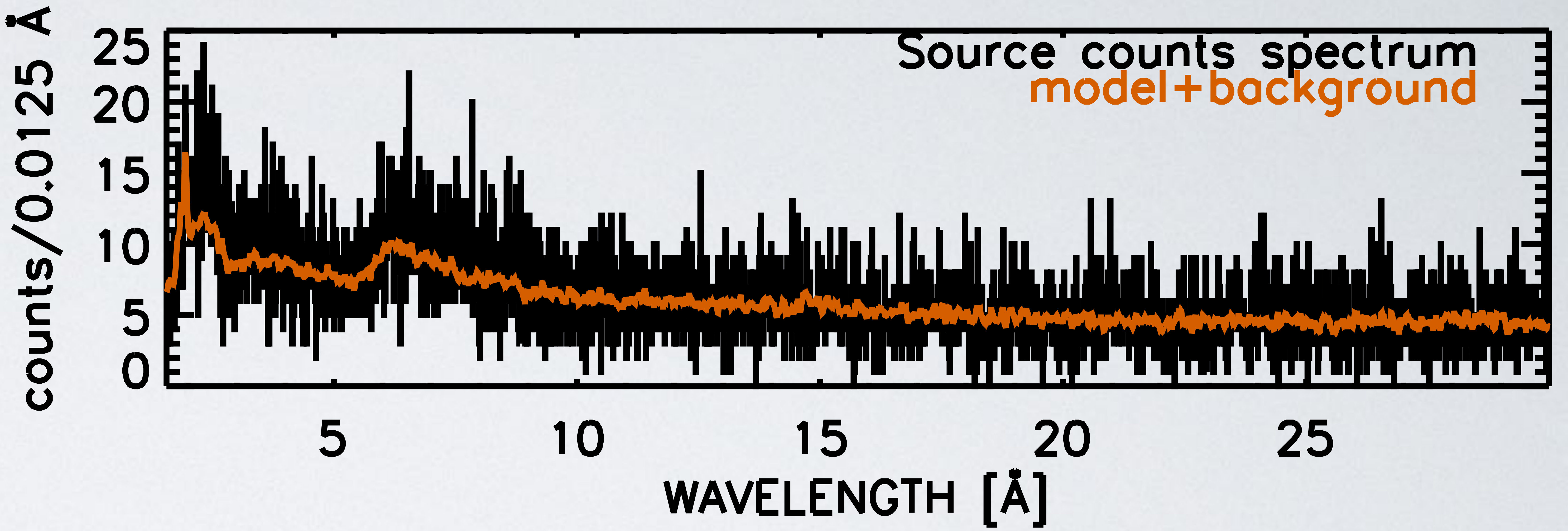


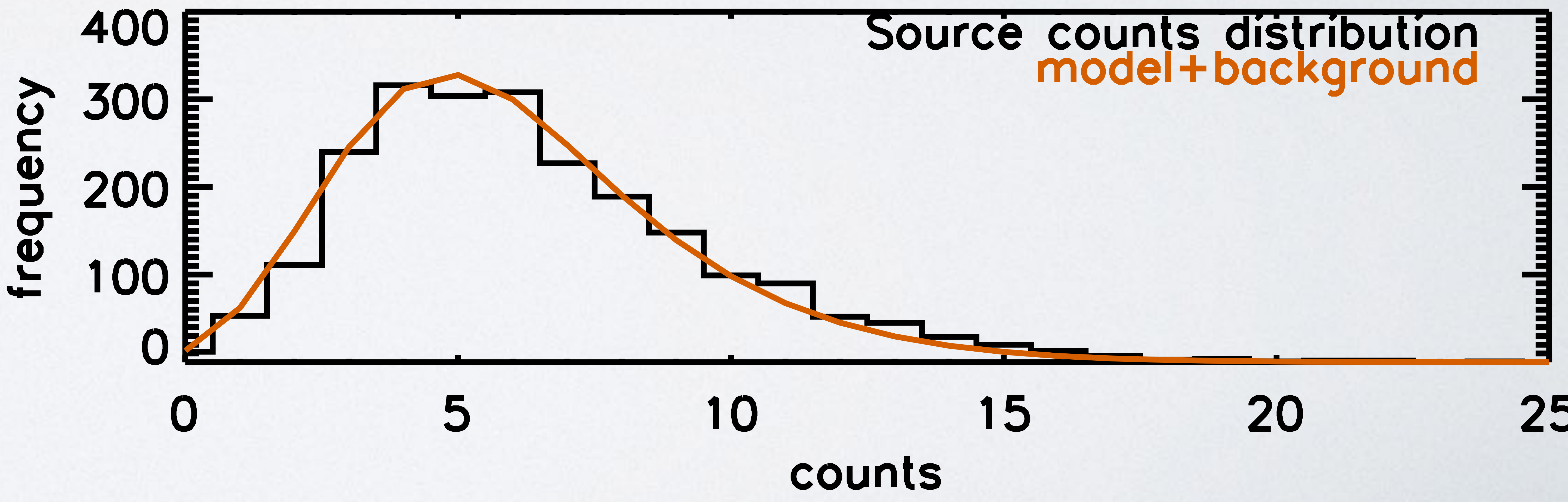




I. WEAK LINES IN HIGH BACKGROUND

- RT Cru is a symbiotic system at 2.5 kpc, with a high mass WD (1.3 $M_\odot$) accreting from a M5III giant
- Exhibits aperiodic flickering, with heavily absorbed hard power-law component, strong lines from Fe XXV, Fe XXVI, and FeKα, and possibly an absorbed soft thermal component
- Observed with Chandra HRC-S/LETG in Nov 2015 for ≈79 ks explicitly to search for lines at longer wavelengths
- **Where are the soft thermal lines?**



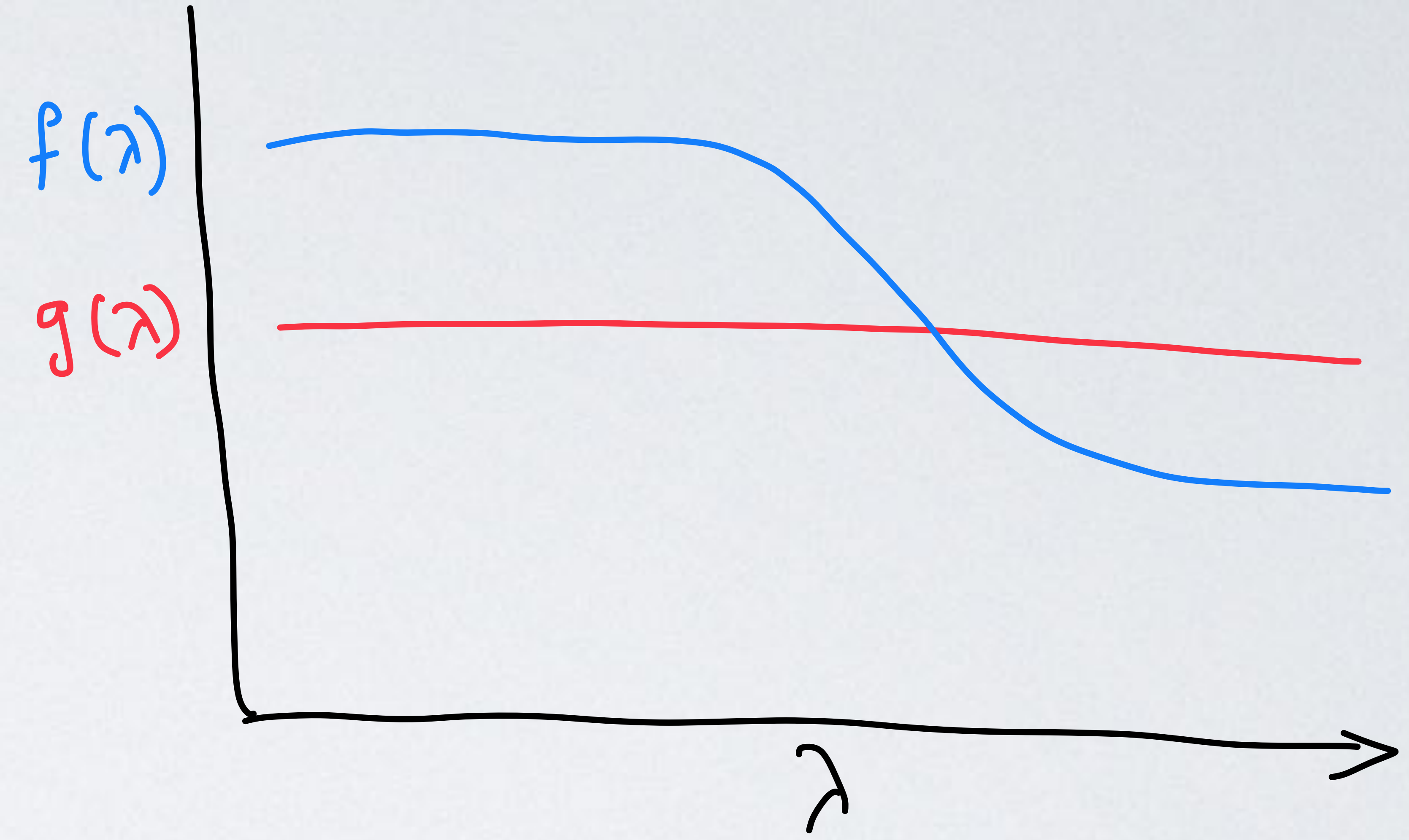


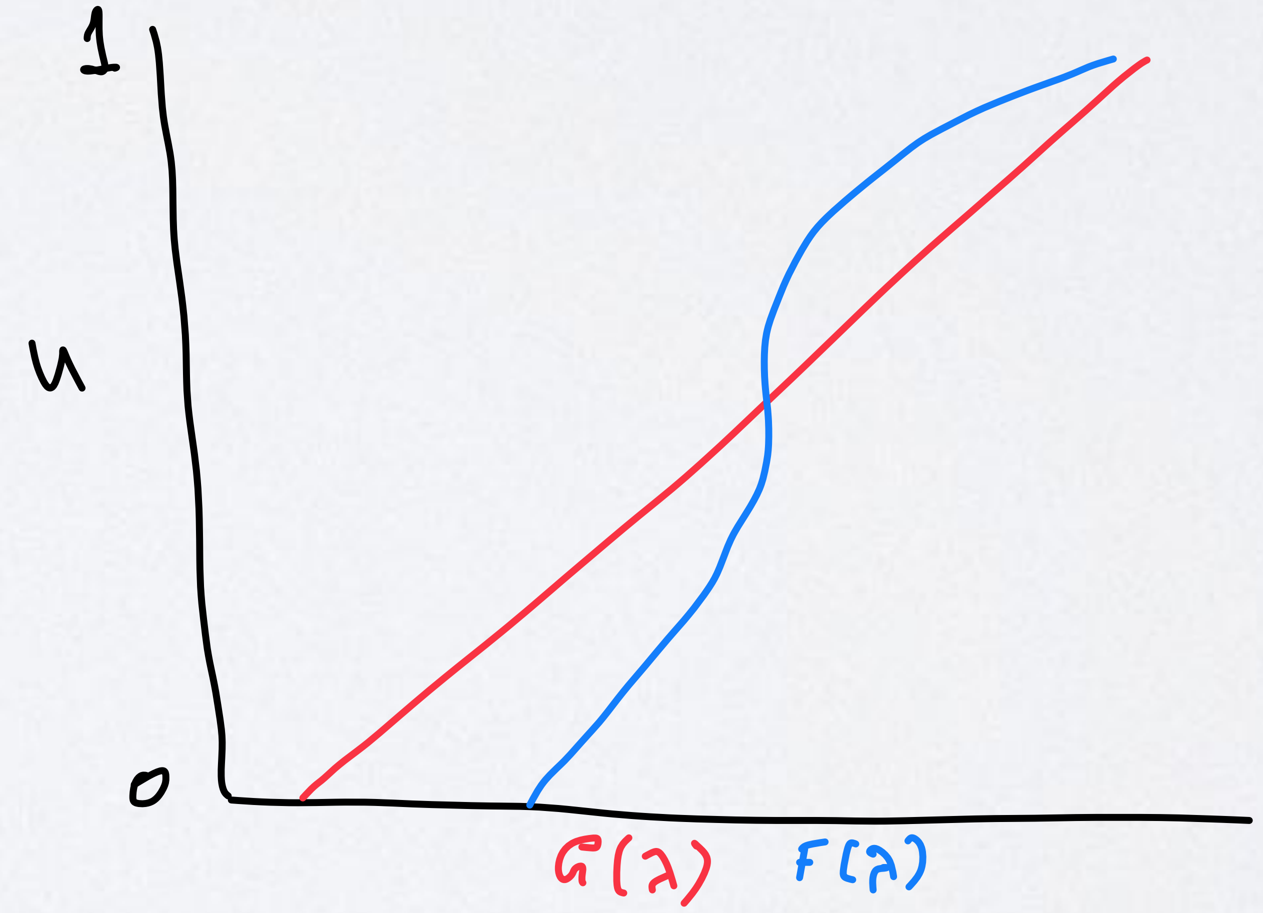
I. HOW WELL DO YOU KNOW YOUR BACKGROUND?

Algeri 2020, PhysRevD, 101, 015003; Zhang et al. 2023, MNRAS, 521, 969

- Suppose you have a model for the background, $g(\lambda)$, but the actual background is $f(\lambda)$
- Trivially, $f(\lambda) = g(\lambda) \cdot [f(\lambda)/g(\lambda)]$
- the ratio of densities can be expressed in quantile form, a comparison density

  $d(u; F,G) = f(G^{-1}(u))/g(G^{-1}(u))$, $u := G(\lambda) \in [0,1]$
- the skew-G density model, a non-parametrically designed parametric modeling of **d(u)**, with orthonormal basis functions, e.g., shifted Legendre polynomials
- number of terms set via a model comparison statistic like BIC



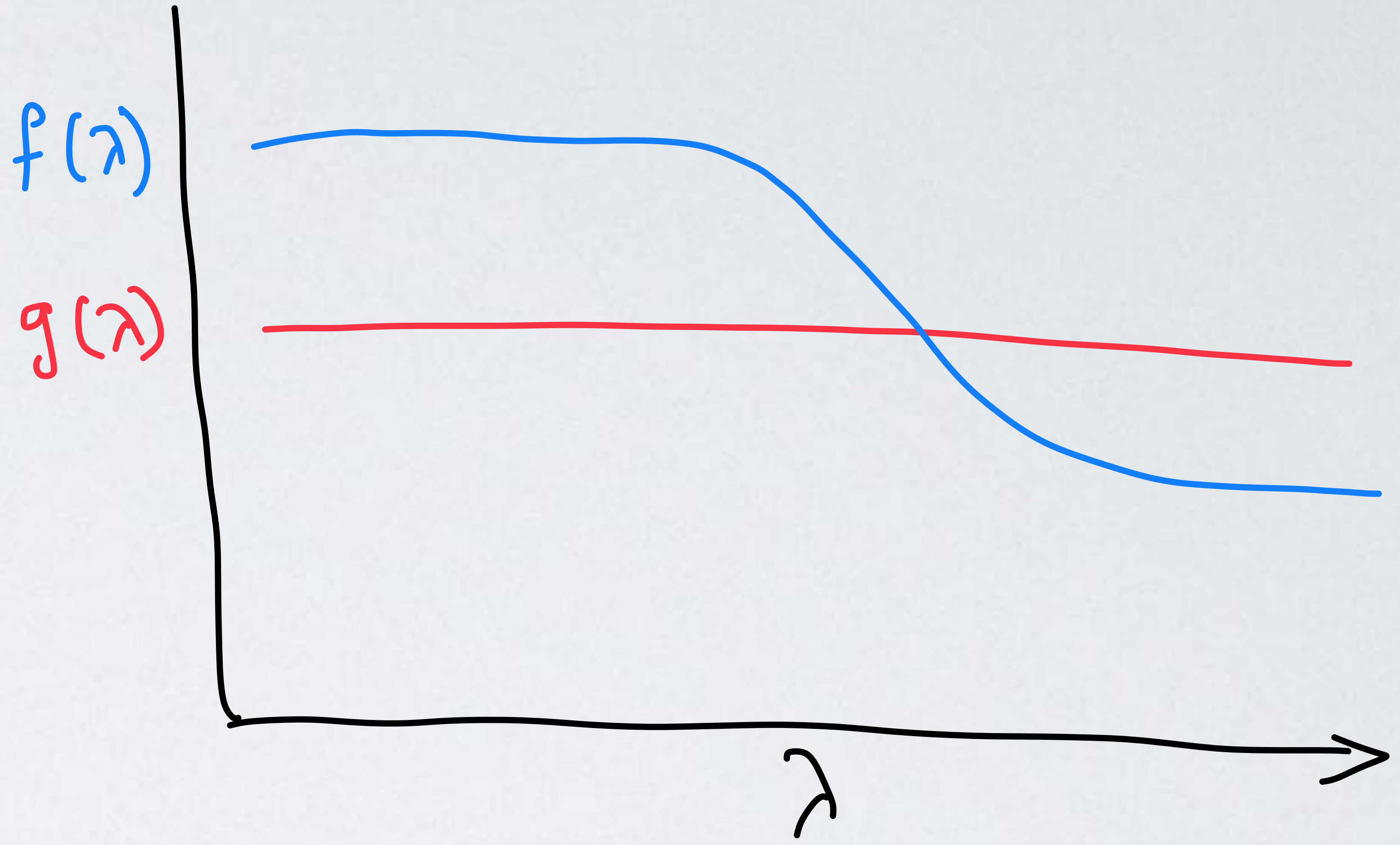


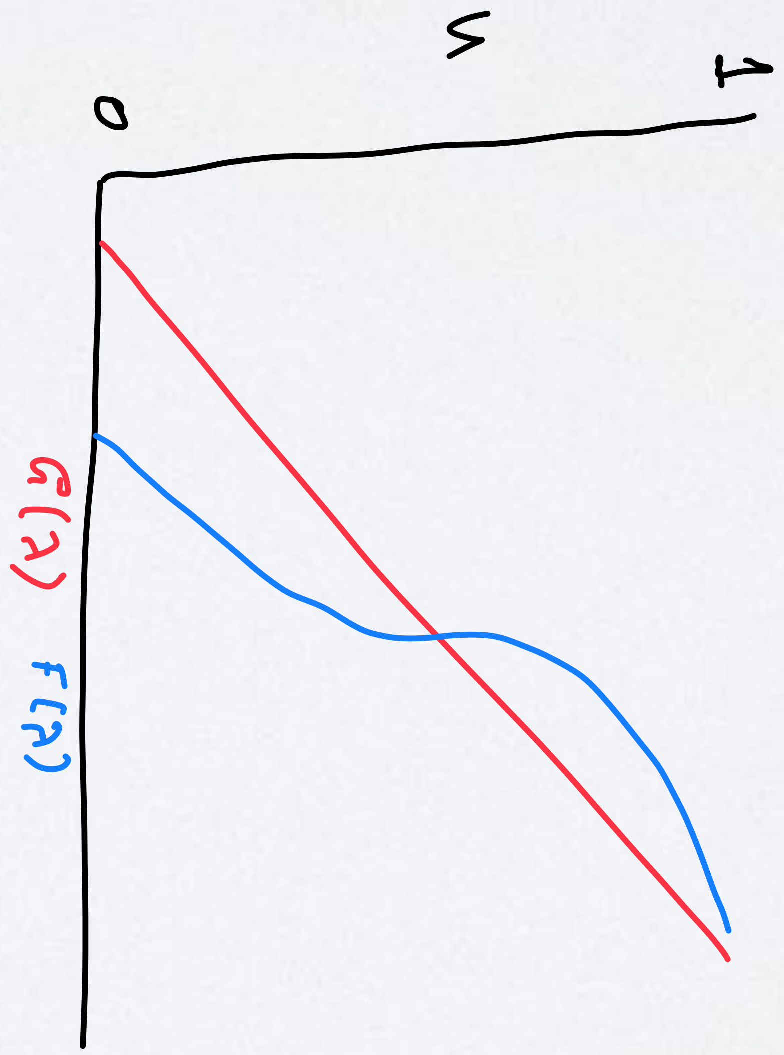
I. HOW WELL DO YOU KNOW YOUR BACKGROUND?

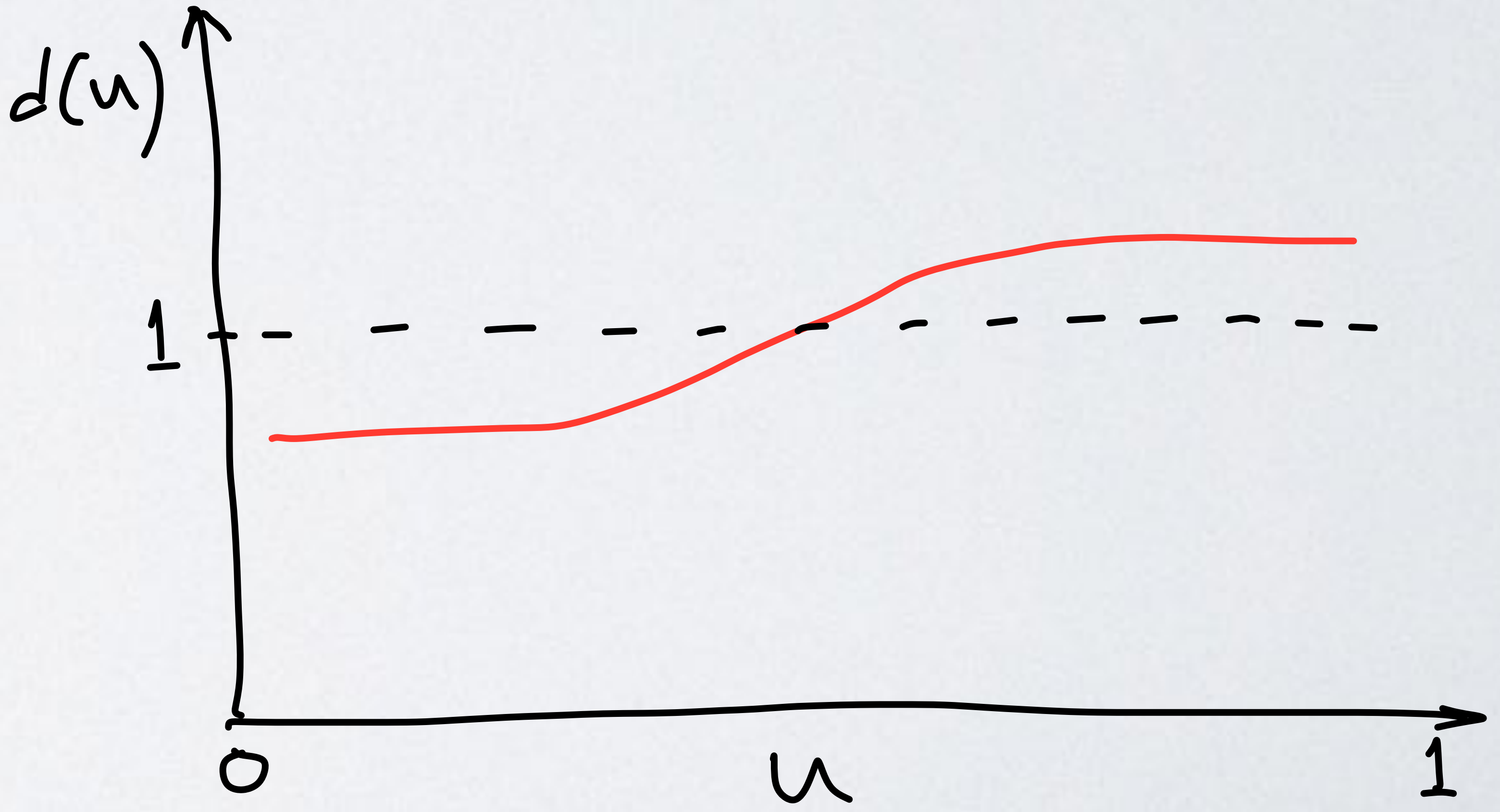
Algeri 2020, PhysRevD, 101, 015003; Zhang et al. 2023, MNRAS, 521, 969

- Suppose you have a model for the background, $g(\lambda)$, but the actual background is $f(\lambda)$
- Trivially, $f(\lambda) = g(\lambda) \cdot [f(\lambda)/g(\lambda)]$
- the ratio of densities can be expressed in quantile form, a comparison density

  $d(u; F,G) = f(G^{-1}(u))/g(G^{-1}(u)), \; u := G(\lambda) \in [0,1]$
- the skew-G density model, a non-parametrically designed parametric modeling of **d(u)**, with orthonormal basis functions, e.g., shifted Legendre polynomials
- number of terms set via a model comparison statistic like BIC







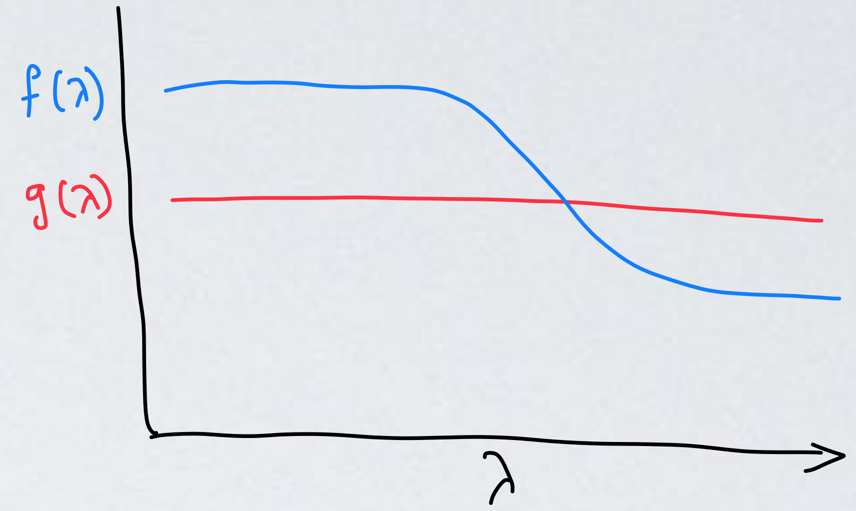
I. HOW WELL DO YOU KNOW YOUR BACKGROUND?

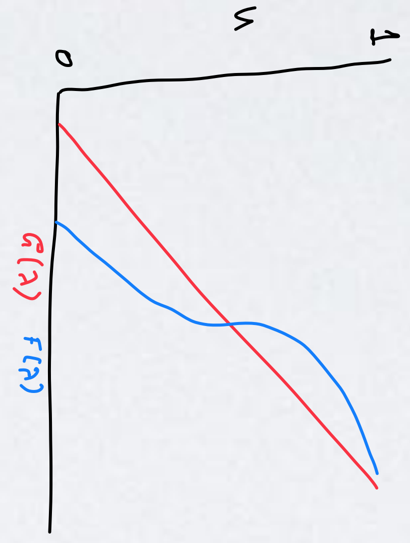
Algeri 2020, PhysRevD, 101, 015003; Zhang et al. 2023, MNRAS, 521, 969

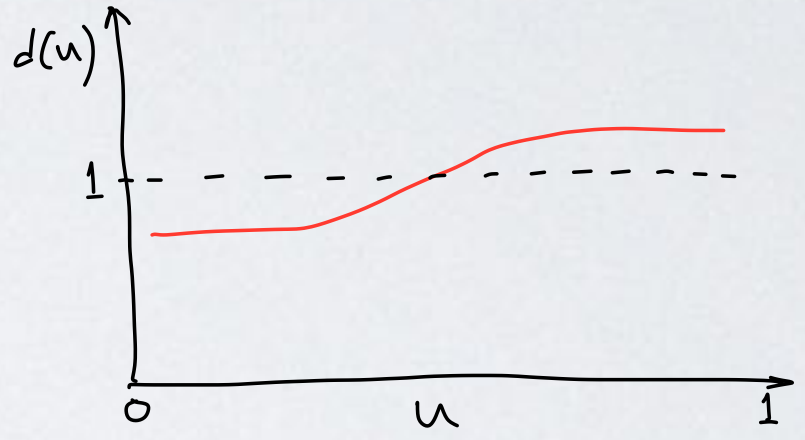
- Suppose you have a model for the background, $g(\lambda)$, but the actual background is $f(\lambda)$
- Trivially, $f(\lambda) = g(\lambda) \cdot [f(\lambda)/g(\lambda)]$
- the ratio of densities can be expressed in quantile form, a comparison density

  $$\mathbf{d(u; F,G)} = f(G^{-1}(u))/g(G^{-1}(u)),\ u := G(\lambda) \in [0,1]$$

- the skew-G density model, a non-parametrically designed parametric modeling of **d(u)**, with orthonormal basis functions, e.g., shifted Legendre polynomials
- number of terms set via a model comparison statistic like BIC





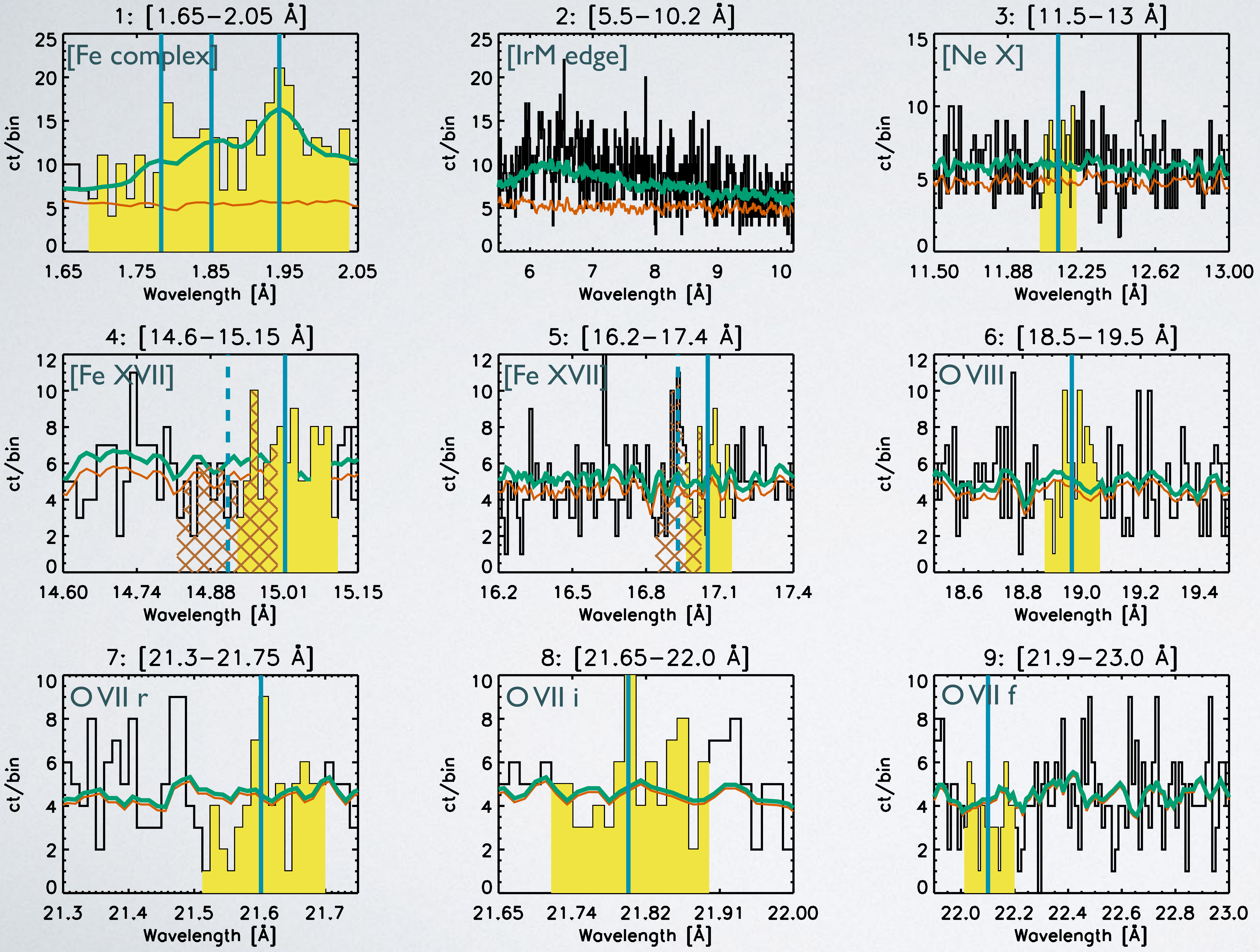


Why do it this way?

1. data-driven measure of complexity in background
2. increase power by discarding information about normalization, work with cumulative distributions
3. easily transferable from background to source data
4. general detection method for arbitrary features

RT CRU UPPER LIMITS VIA POWER

Zhang et al. 2023, MNRAS



Some of the expected strong lines from thermal emission.

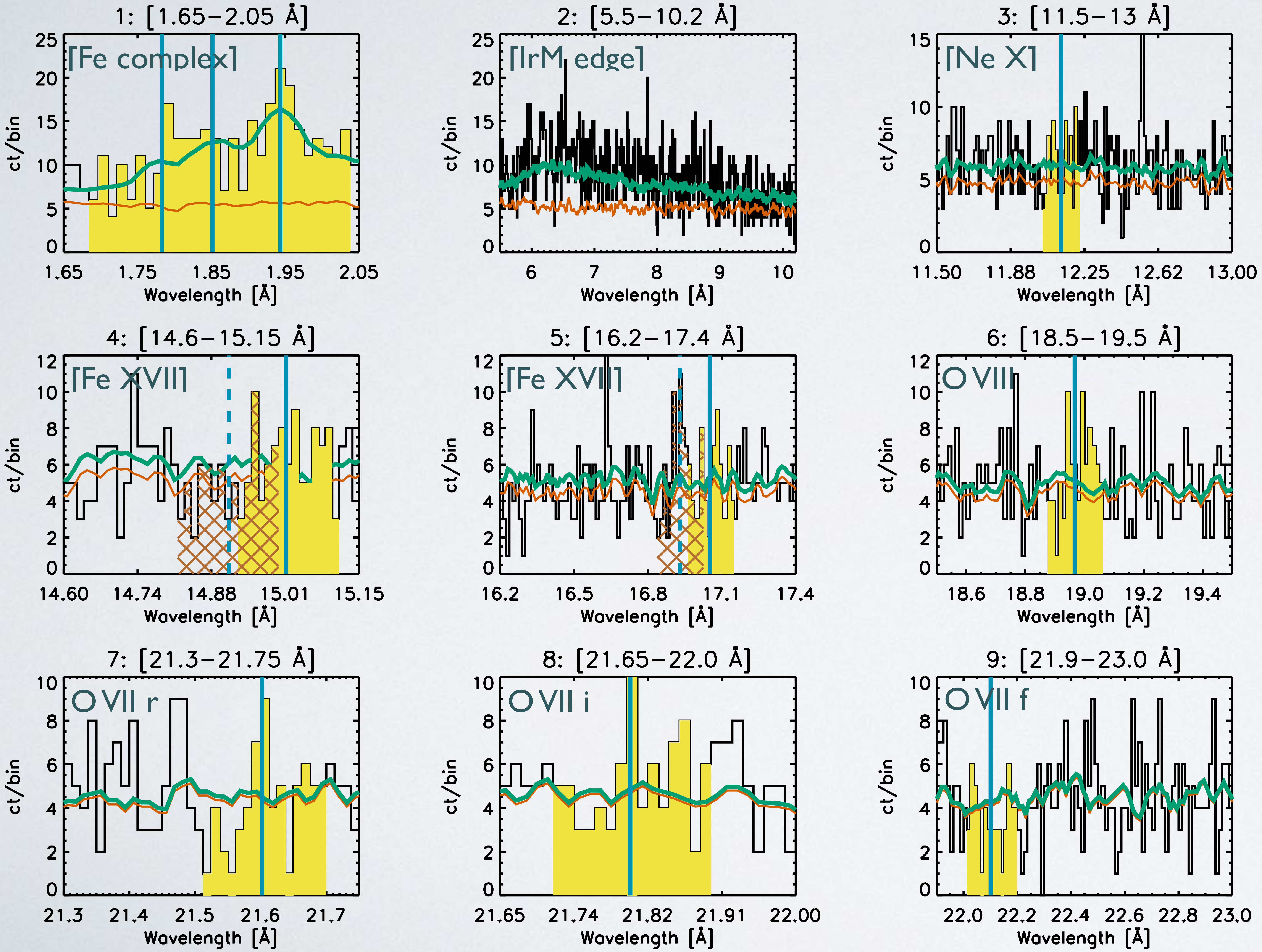
Locations of lines are marked by vertical lines and line spreads by yellow shades.

Fe lines, IrM edge, and unidentified line at 16.93 Å are detected.

The rest require upper limits to be set.

RT CRU UPPER LIMITS VIA POWER

Zhang et al. 2023, MNRAS



Regions of interest (W_r)	m	Bonferroni (Sidak)	K (Sidak)	Naive (Sidak)
W_1	3	0.0001 (0.0011)	0.0071 (0.0397)	0.0045 (0.0621)
W_2	3	1.0816e-18 (1.0817e-17)	2.7907e-15 (2.9976e-14)	3.3306e-15 (2.4980e-14)
W_3	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
W_4	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
W_5	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
W_6	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
W_7	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
W_8	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
W_9	0	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)

Testing for difference from background

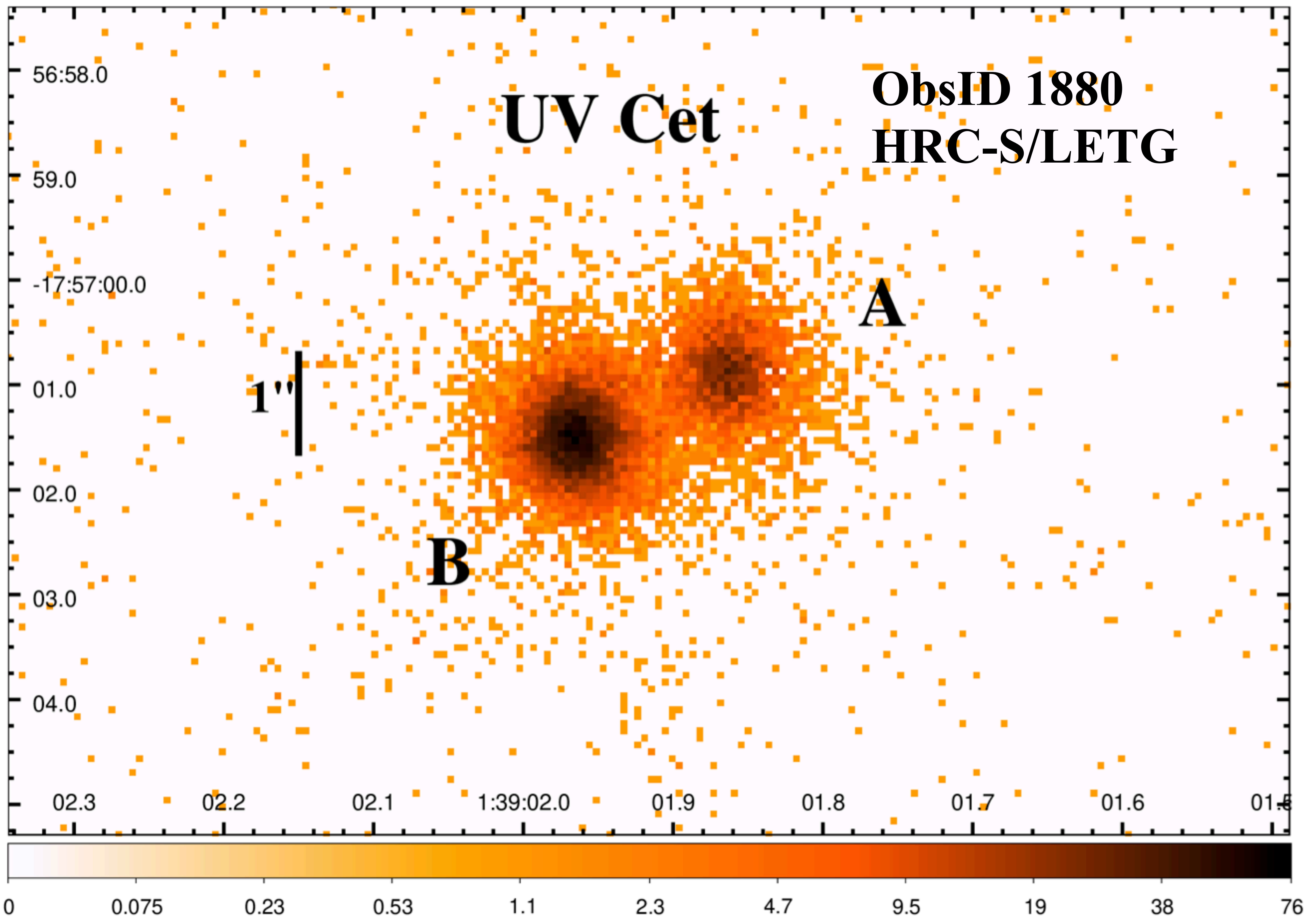
Testing for lines at nominal locations

Regions of interest (W_r)	Local p-values	Sidak's correction
W_3	0.4810	0.9899
W_4	0.1143	0.5724
W_5	0.3247	0.9359
W_6	0.0385	0.2402
W_7	0.2612	0.8799
W_8	0.5000	0.9922
W_9	0.5000	0.9922

Setting upper limits to lines

Regions (W_r)	50% upper limits via LRT		90% upper limits via LRT	
	Local	Sidak adjusted	Local	Sidak adjusted
W_3	29.93	39.42	48.91	53.29
W_4	20.00	26.43	32.36	39.52
W_5	24.02	30.14	35.32	43.80
W_6	22.62	28.08	34.71	39.39
W_7	17.90	24.17	29.71	35.98
W_8	17.84	24.80	30.30	36.25
W_9	37.83	21.87	63.57	76.83

II. DISAMBIGUATE OVERLAP OF UV CET A & B



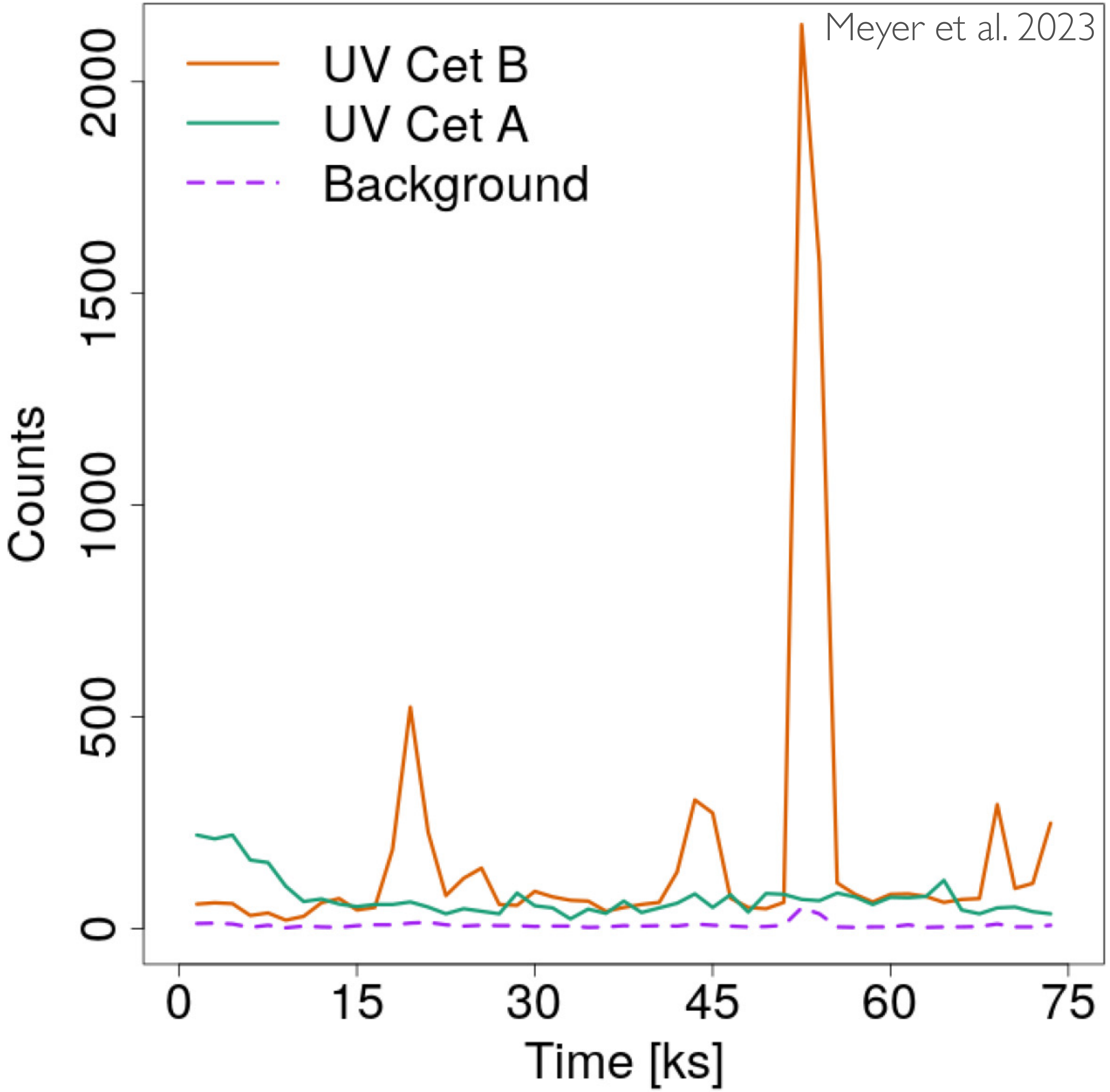
UV Cet AB ≡ GJ 65 AB

Flaring active dM binary
M5.5V+M6V
≈2.7 pc

Observed for 75 ks in Nov 2001 with HRC-S/LETG

Massive >100× flare on UV Cet B

II. EBASCS ON 0TH ORDER



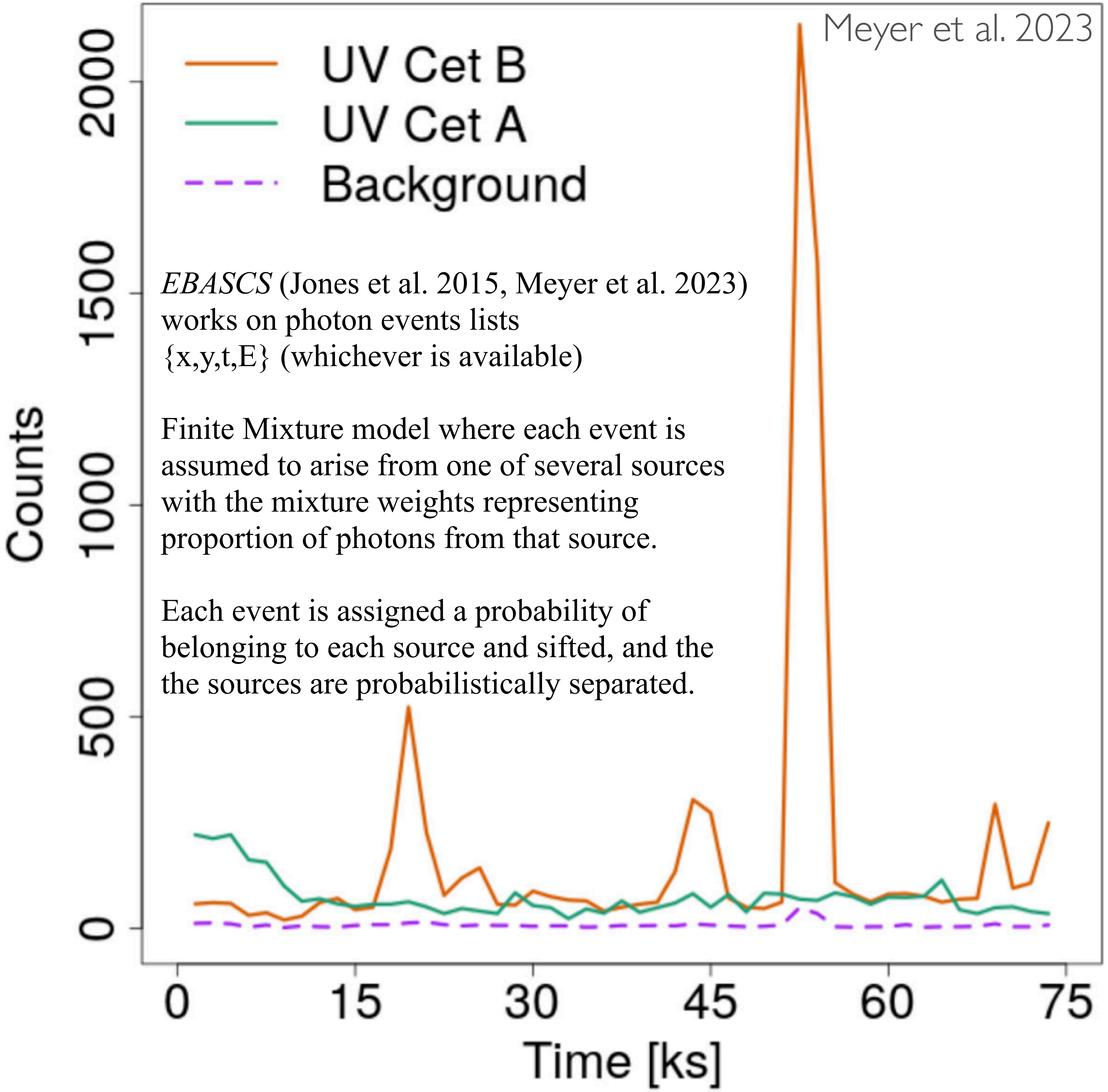
$L_{\text{Xpeak}} \approx 2{\cdot}10^{29}$ erg $s^{-1} \approx$ X1000
$L_{\text{Xmin}} \approx 10^{27}$ erg $s^{-1} \approx 200\times$ increase
(Audard et al. 2003)

EBASCS (Jones et al. 2015, Meyer et al. 2023) works on photon events lists {x,y,t,E} (whichever is available)

Finite Mixture model where each event is assumed to arise from one of several sources with the mixture weights representing proportion of photons from that source.

Each event is assigned a probability of belonging to each source and sifted, and the the sources are probabilistically separated.

II. EBASCS ON 0TH ORDER



EBASCS (Jones et al. 2015, Meyer et al. 2023) works on photon events lists {x,y,t,E} (whichever is available)

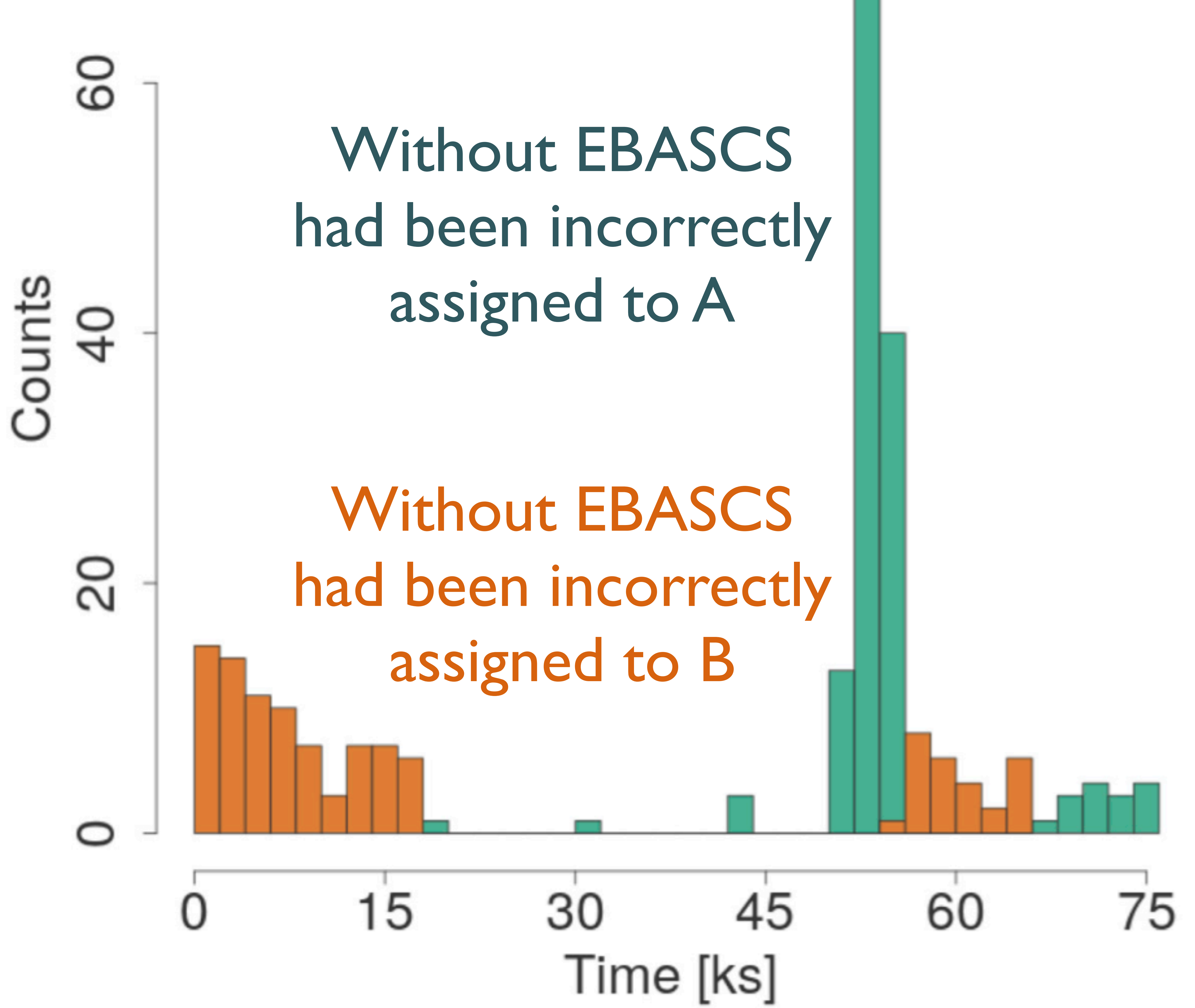
Finite Mixture model where each event is assumed to arise from one of several sources with the mixture weights representing proportion of photons from that source.

Each event is assigned a probability of belonging to each source and sifted, and the the sources are probabilistically separated.

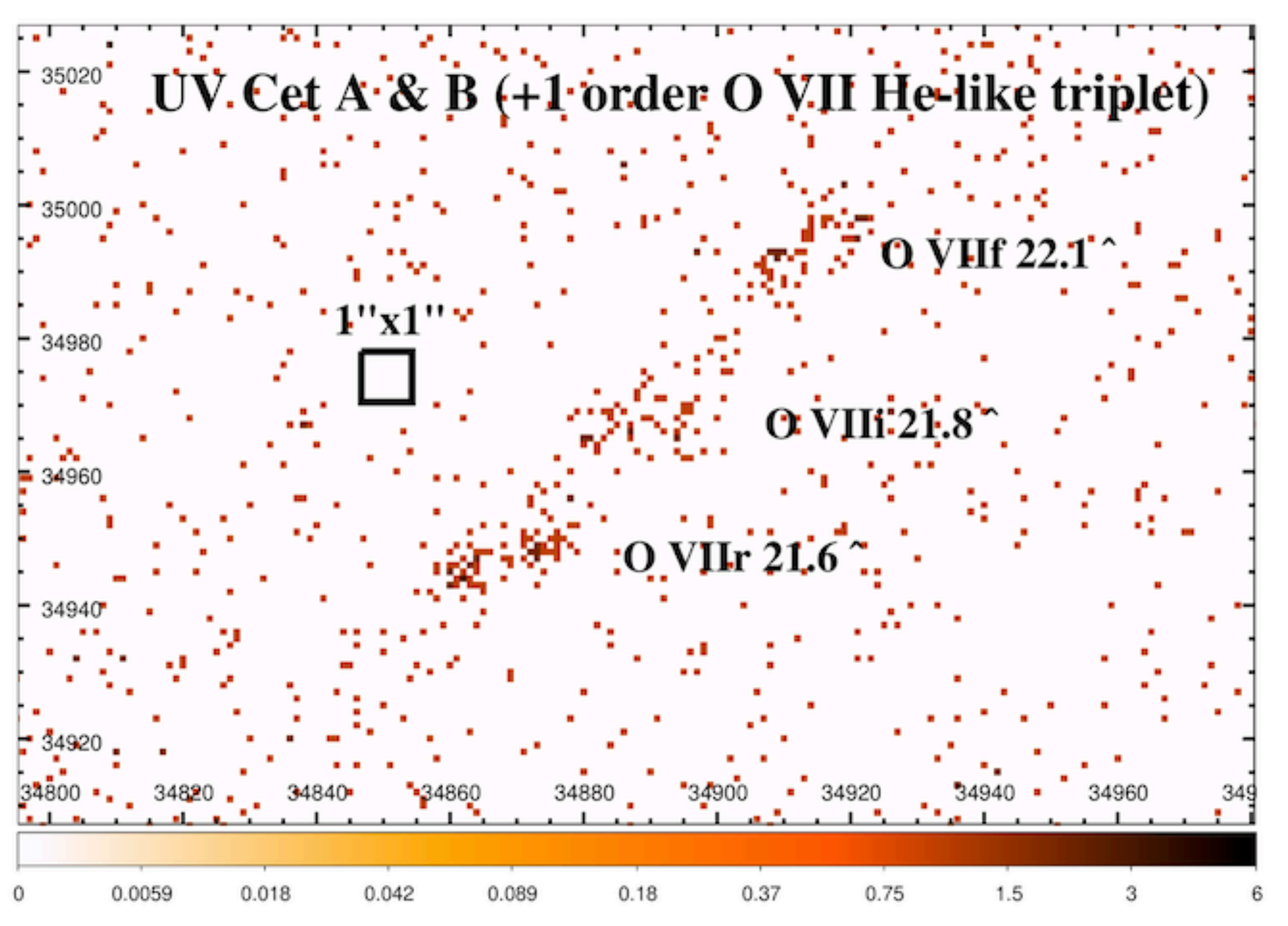
$L_{\mathrm{Xpeak}} \approx 2{\cdot}10^{29}$ erg s^{-1} $\approx$ X1000

$L_{\mathrm{Xmin}} \approx 10^{27}$ erg s^{-1} $\approx$200× increase

(Audard et al. 2003)



II. DISAMBIGUATE OVERLAP OF UV CET A & B O VII TRIPLET

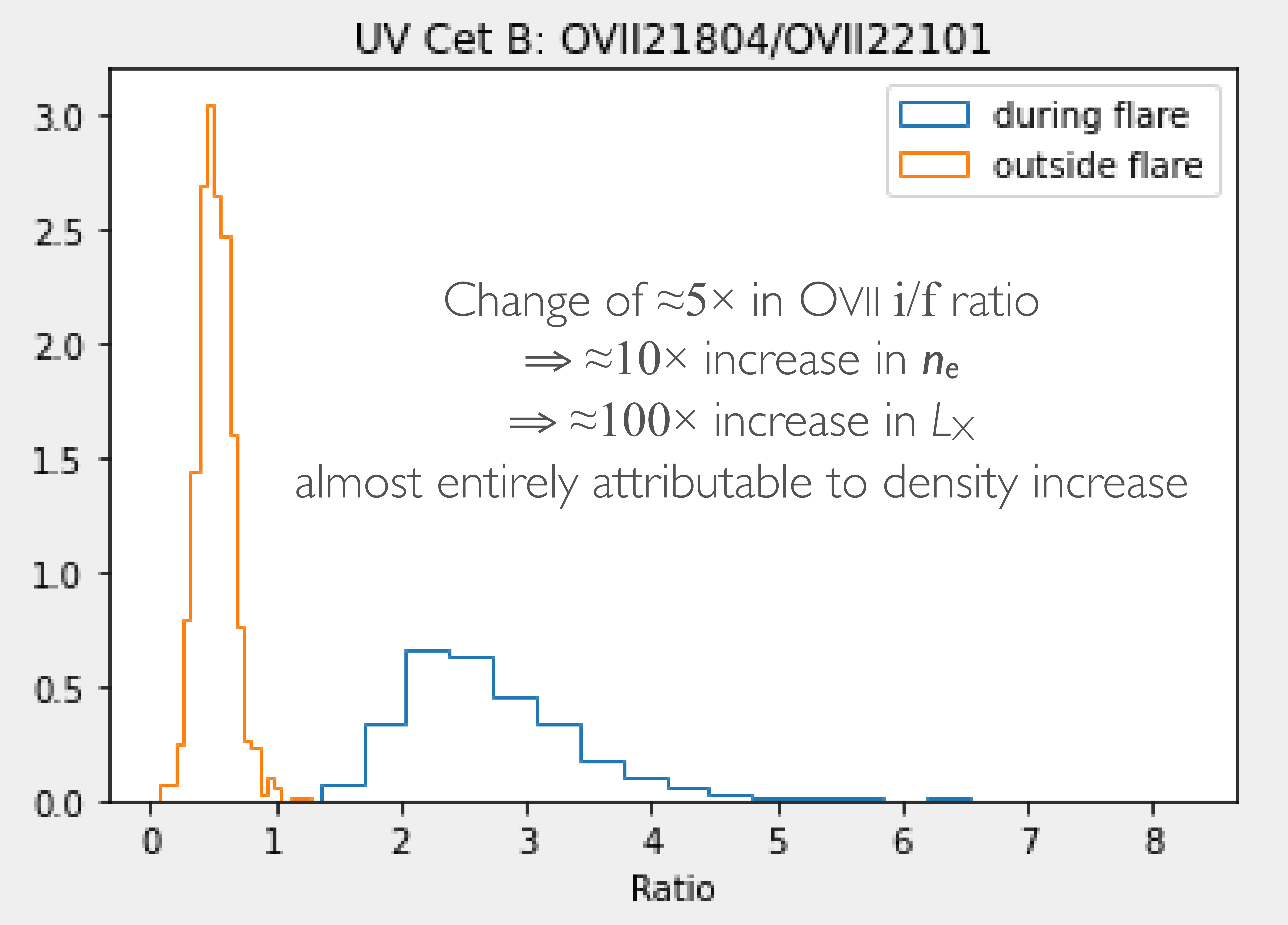


Model the OVII triplet

Propagate the relative positions of A & B from 0th order to expected locations of dispersed O VII triplet

Fit weights relative to O VIIr B

II. EBASCS ON DISPERSED OVII TRIPLET



SUMMARY

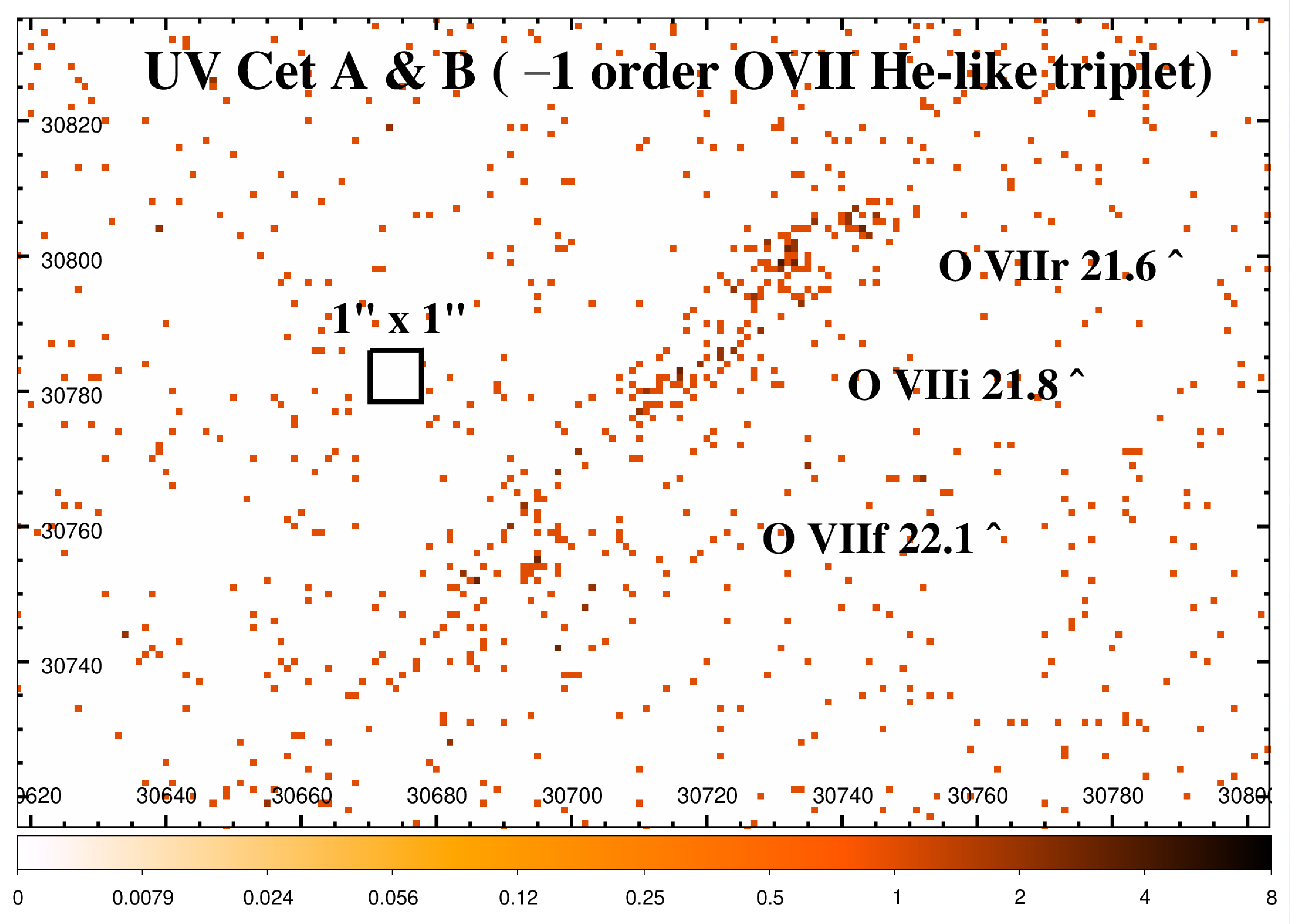
I. Weak lines in high background

- ✦ New method to discern deviations in background model and detect presence of weak source features, accounting for multiple tests
- ✦ Applied to look for soft lines in RT Cru — none of the usual suspects are *detectable* with current instrumentation, set upper limits on line fluxes
- ✦ What is that line at 16.93 Å?

II. Disentangling overlapping lines from contaminating companions

- ✦ Probabilistically sift photons in overlapping sources using spatial, spectral, and temporal differences
- ✦ Applied to UV Cet O VII density sensitive lines, demonstrates rapid increase of density during flare
- ✦ Developed for grating spectroscopy, will also work for calorimeter detectors

II. DISAMBIGUATE OVERLAP OF UV CET A & B



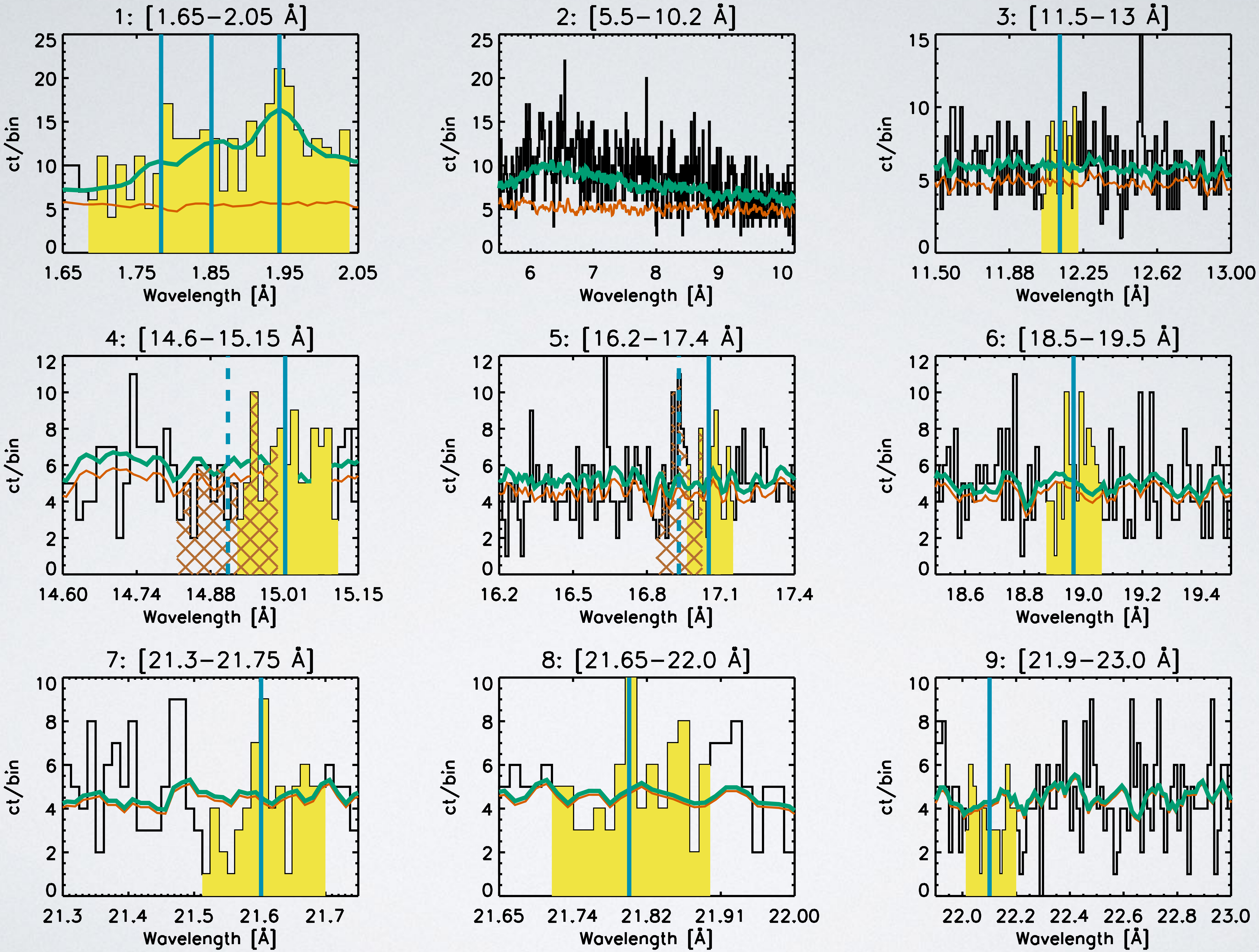
Model the OVII triplet

Propagate the relative positions of A & B from 0th order to expected locations of dispersed O VII triplet

Fit weights relative to O VIIr B

[DIGRESSION] UPPER LIMITS: A RANT

1. An Upper Limit is *not* the upper bound of the uncertainty interval of a flux estimate
    a. An uncertainty interval is not unique — a 68% uncertainty interval on flux can be *anything* between $[0,q_{68}]$ to $[q_{32},\infty]$, even $[q_{16},q_{84}]$
2. An upper *limit* is how bright a source could be before it will be definitely detected, or how faint should it be for it to be definitely *not* detected (Kashyap et al. 2010, ApJ 719, 900)
3. You set an upper limit based on the *process of detection*, not based on how many counts are observed for the source, because then you have an *estimate* of the flux
4. It requires a measure of the False Negative, or Statistical Power
5. See (1)


1: [1.65–2.05 Å]
2: [5.5–10.2 Å]
3: [11.5–13 Å]
4: [14.6–15.15 Å]
5: [16.2–17.4 Å]
6: [18.5–19.5 Å]
7: [21.3–21.75 Å]
8: [21.65–22.0 Å]
9: [21.9–23.0 Å]
ct/bin
Wavelength [Å]