Spectro-Perfectionism in SDSS-III

Adam S. Bolton

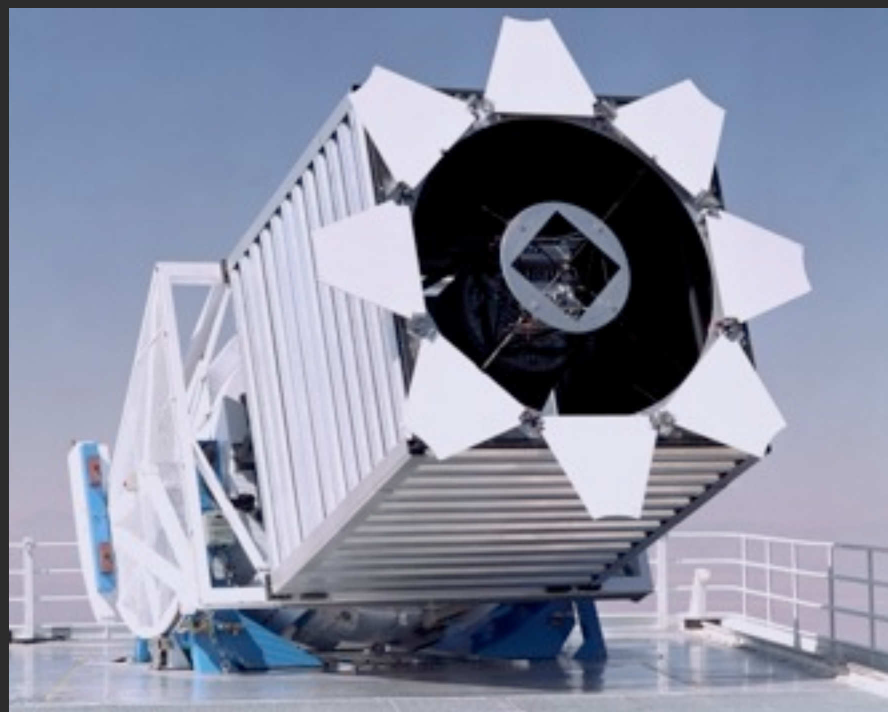
Department of Physics & Astronomy

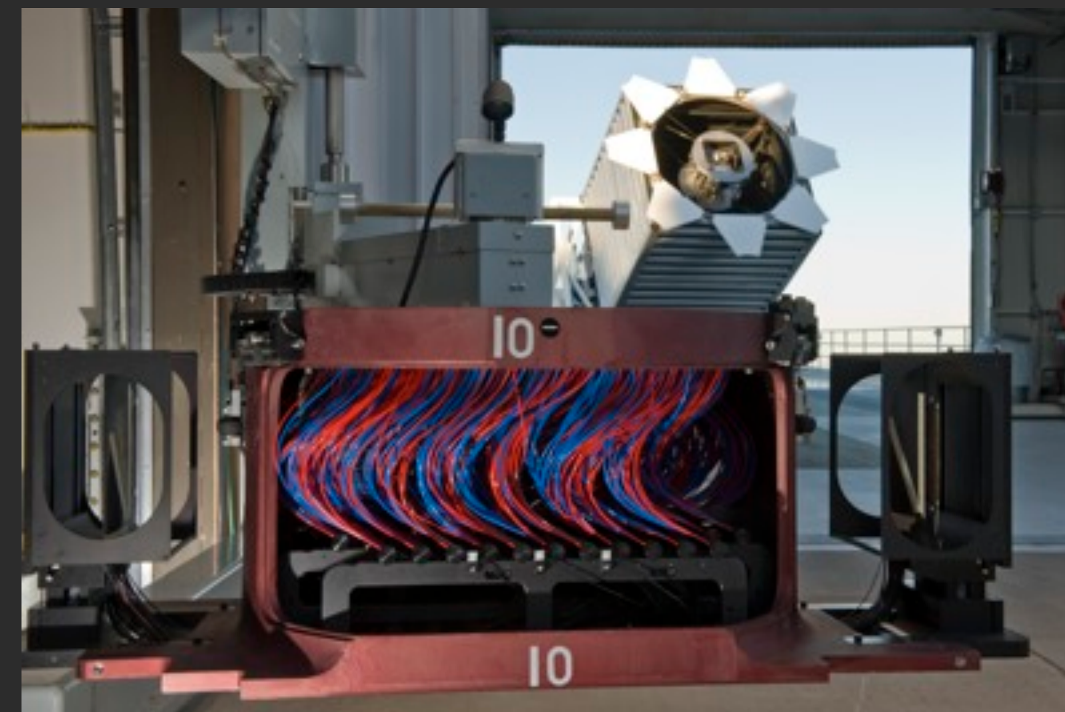
The University of Utah

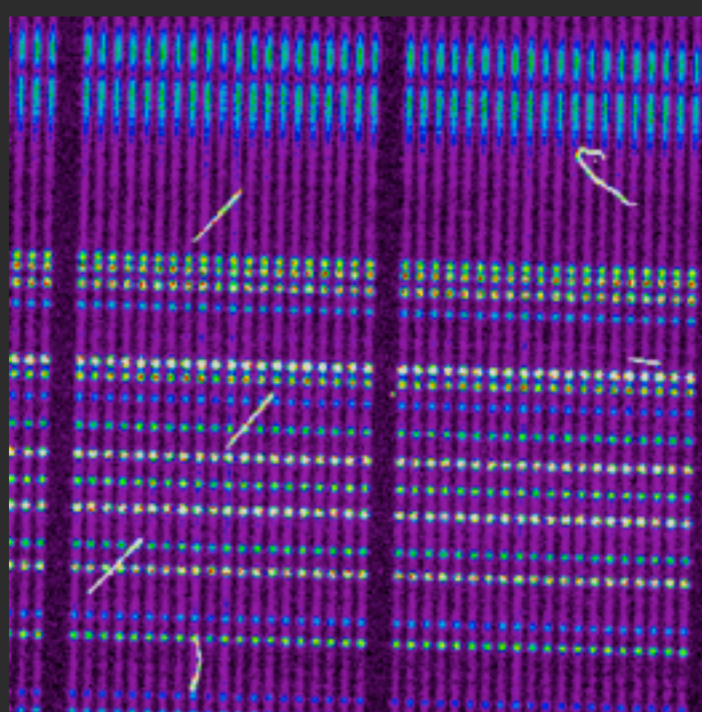
ADASS XXI - Paris - 2011-11-09

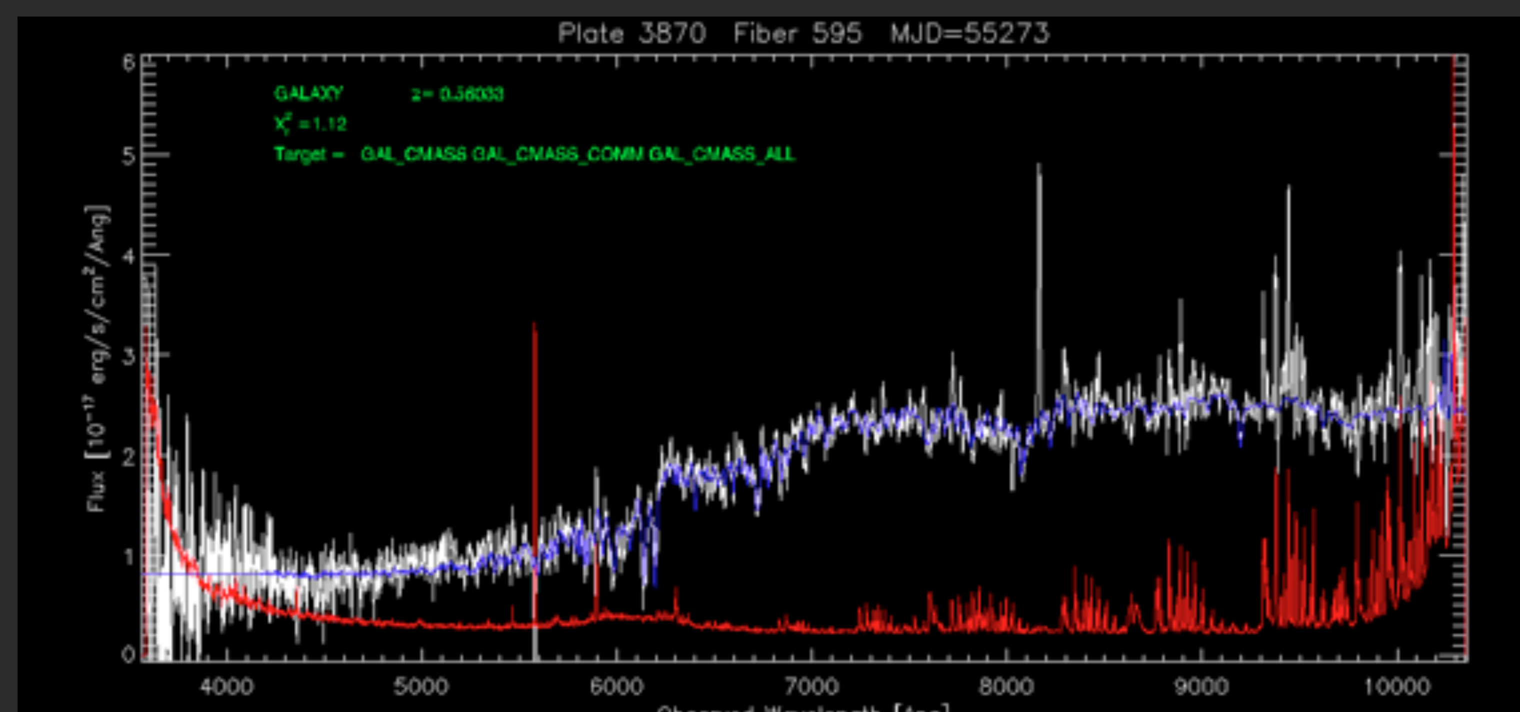
What is SDSS-III?

Eisenstein et al. 2011









What is SDSS-III?

Eisenstein et al. 2011

BOSS: The Baryon Oscillation Spectroscopic Survey

- One of the four SDSS-III surveys
- 2009-2013 spectroscopic operations
- Redshifts of 1.5 million galaxies to $z = 0.7$
- 160k quasars for Lyman-α forest
- Measurement of baryon acoustic feature vs. z
- Constrain parameters of "dark energy"
- Largest spectro data set for massive galaxy evolution

What is...

Spectro-Perfectionism
a.k.a.
2D PSF Extraction
a.k.a.
the Bolton & Schlegel algorithm
?

(Bolton & Schlegel 2010, PASP, 122, 248)

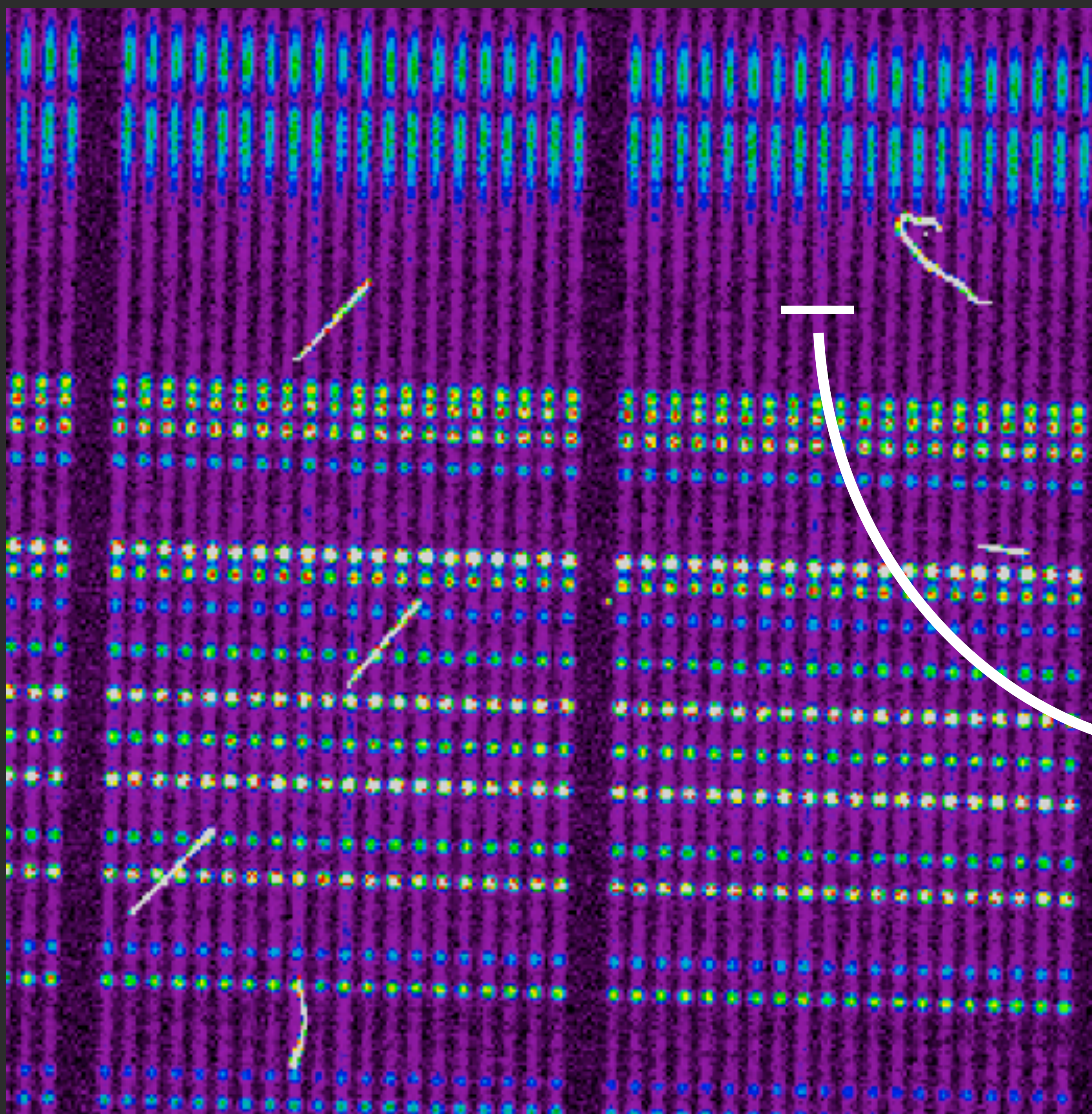
What is...

Spectro-Perfectionism
a.k.a.
2D PSF Extraction
a.k.a.
the Bolton & Schlegel algorithm
?

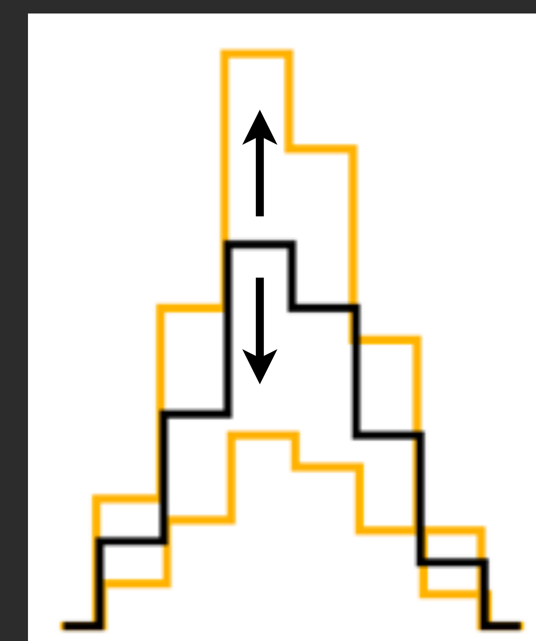
(Bolton & Schlegel 2010, PASP, 122, 248)

Spectroscopic extraction via mathematically correct forward modeling of the raw data via the 2D spectrograph point-spread function (PSF).

Doesn’t “optimal extraction” do this?

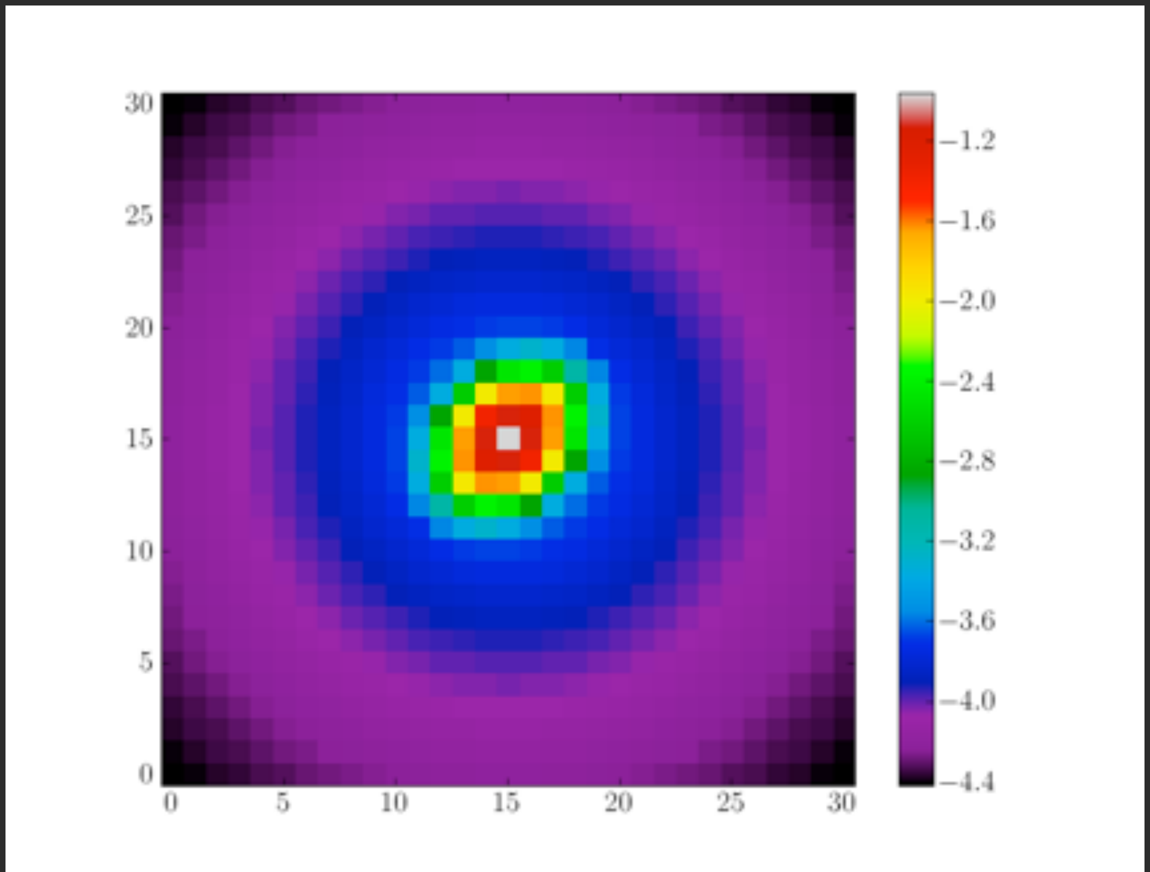


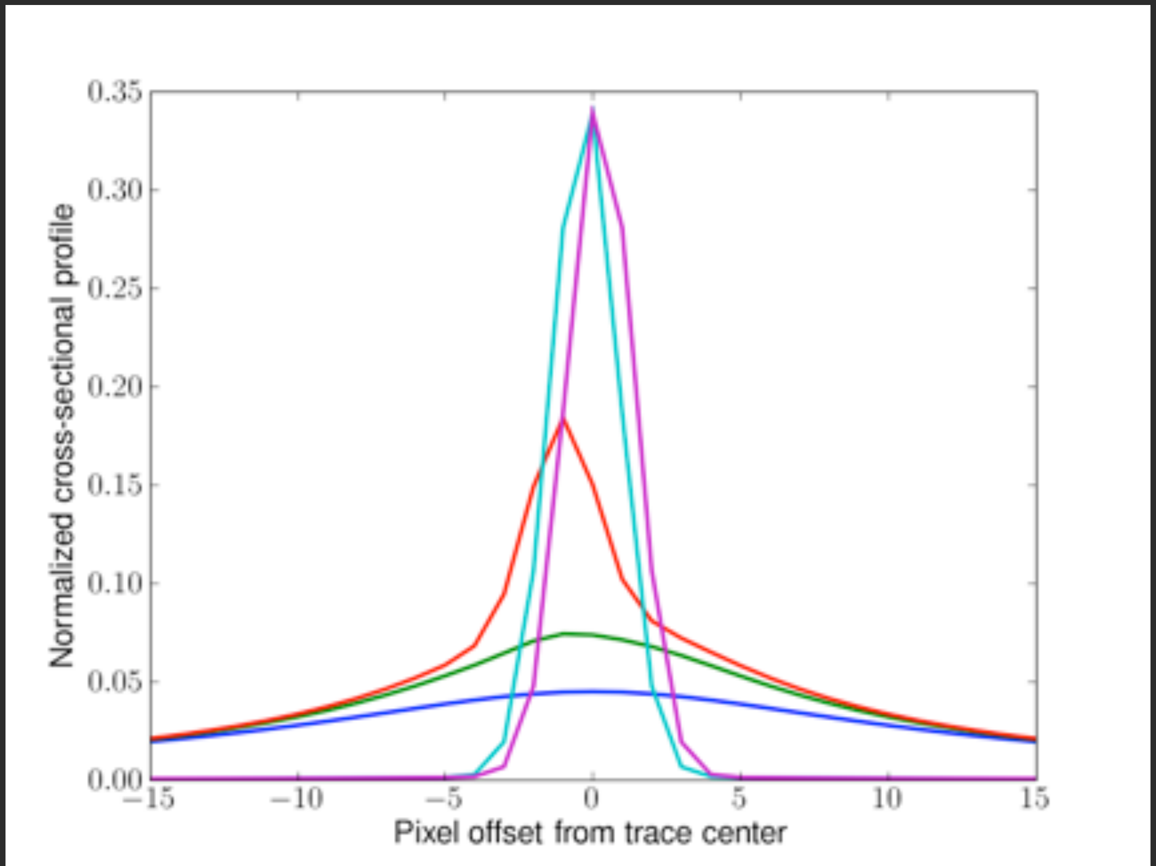
Hewett et al. 1985;
Horne 1986



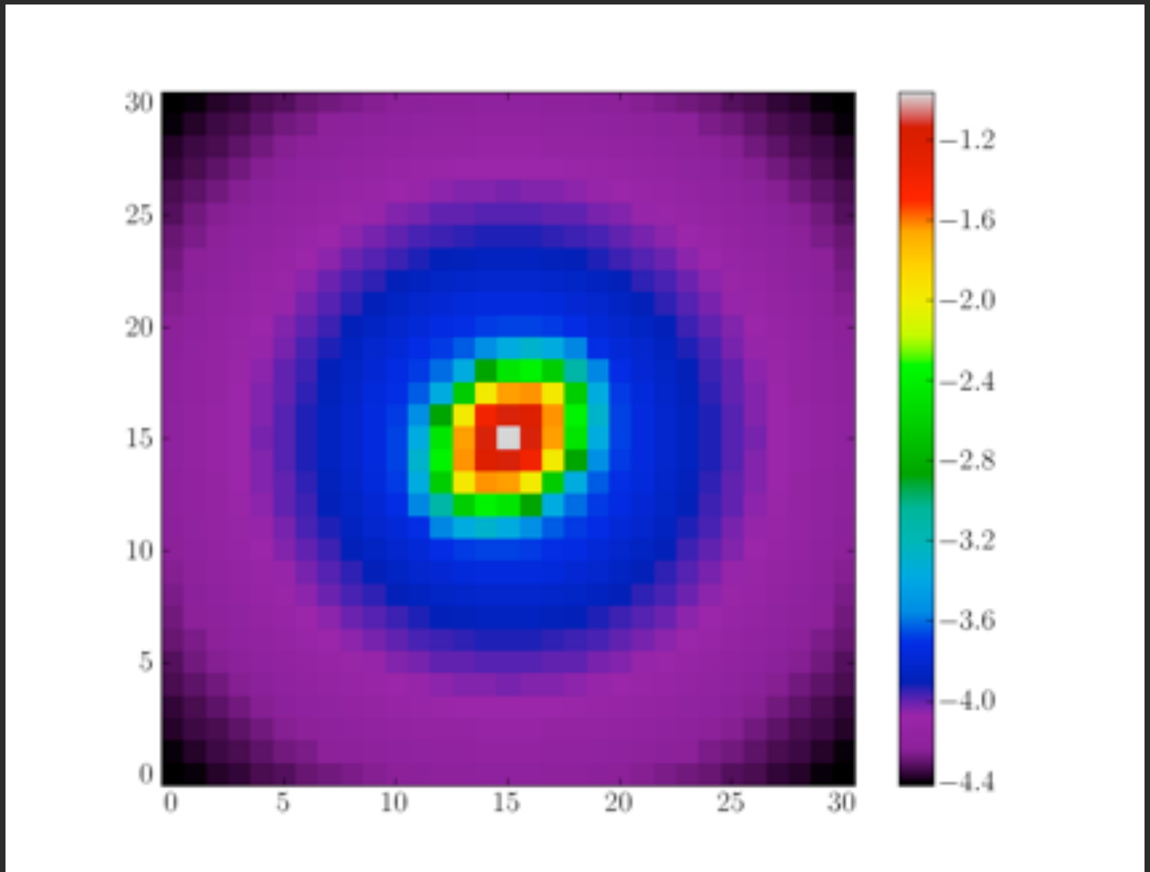
- Determine cross-sec’n
- Weighted amplitude fit
- Call that your spectrum

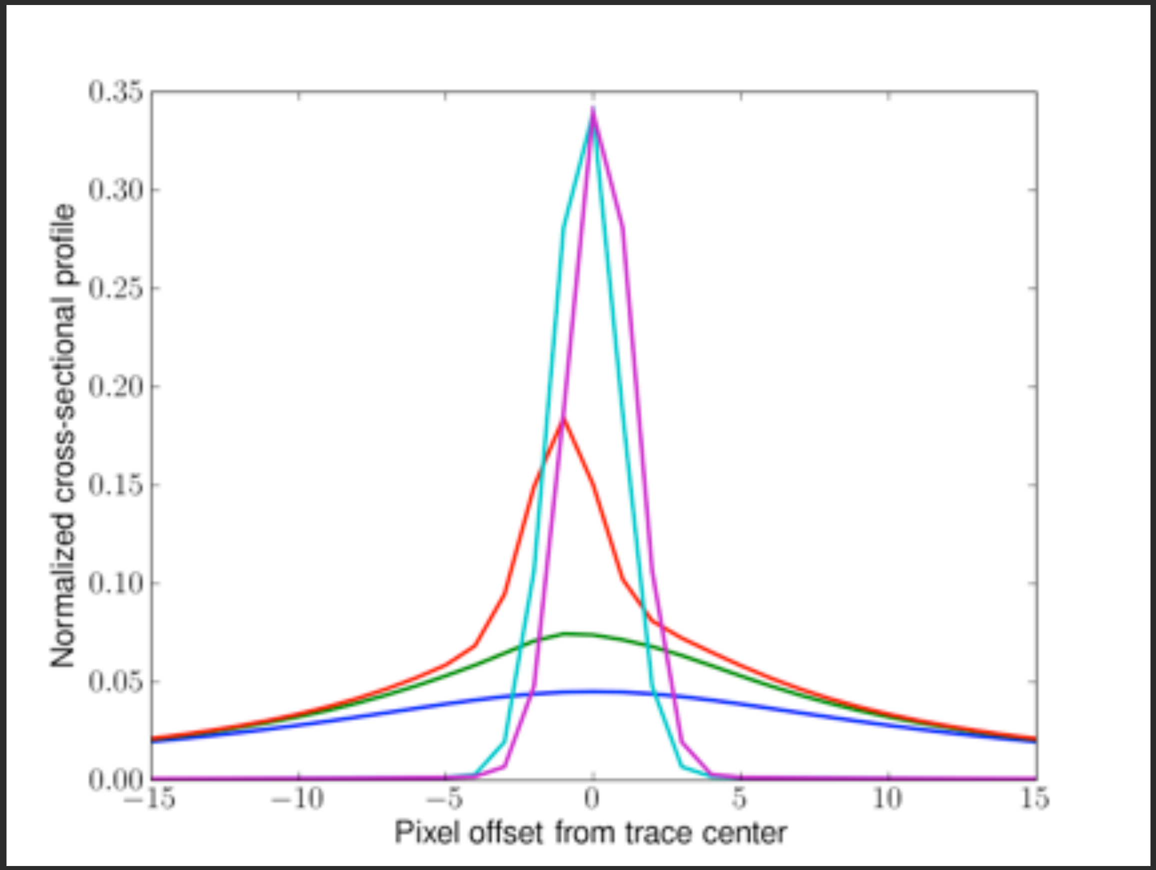
How do you extract an emission line?





How do you extract an emission line?

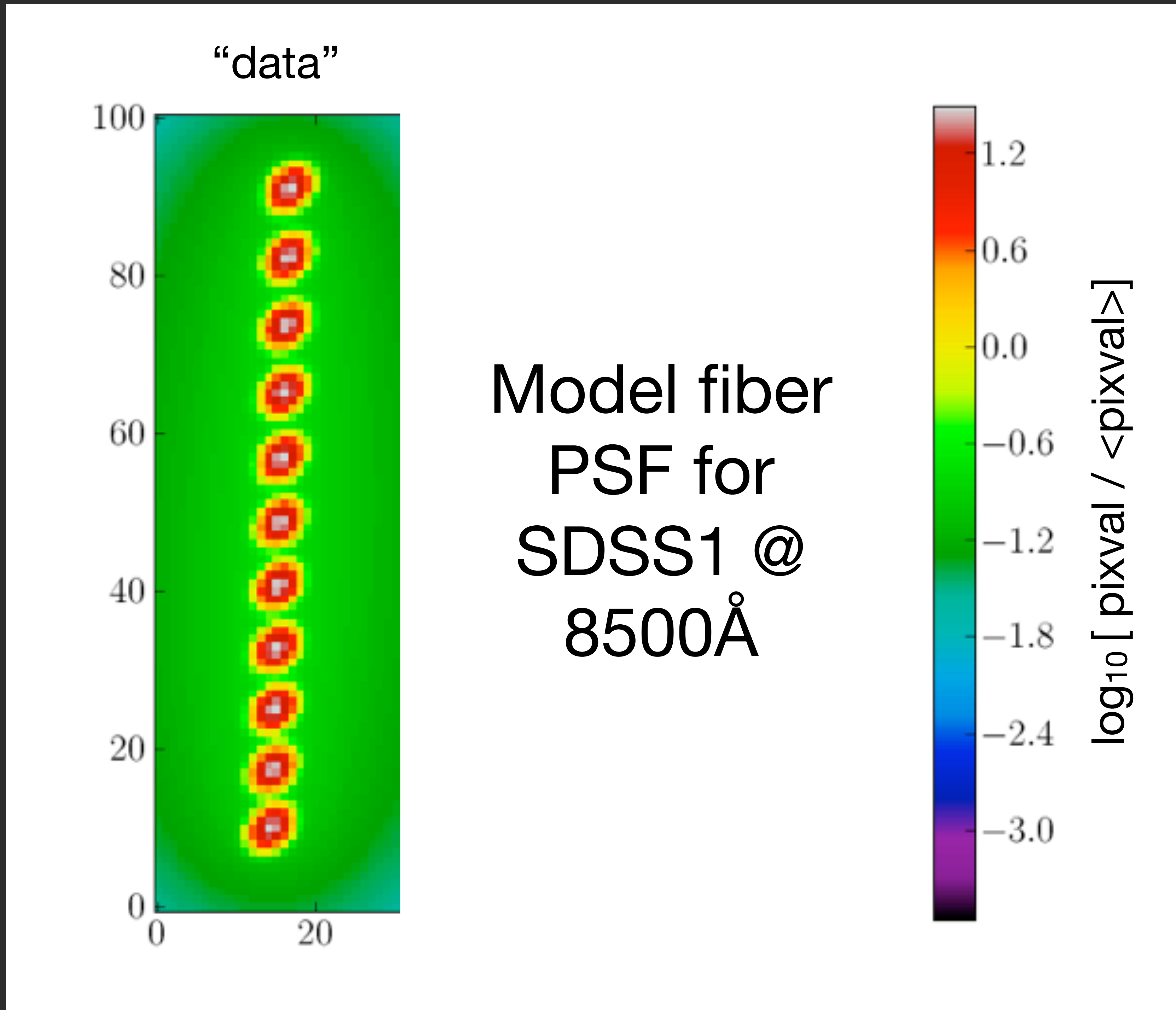




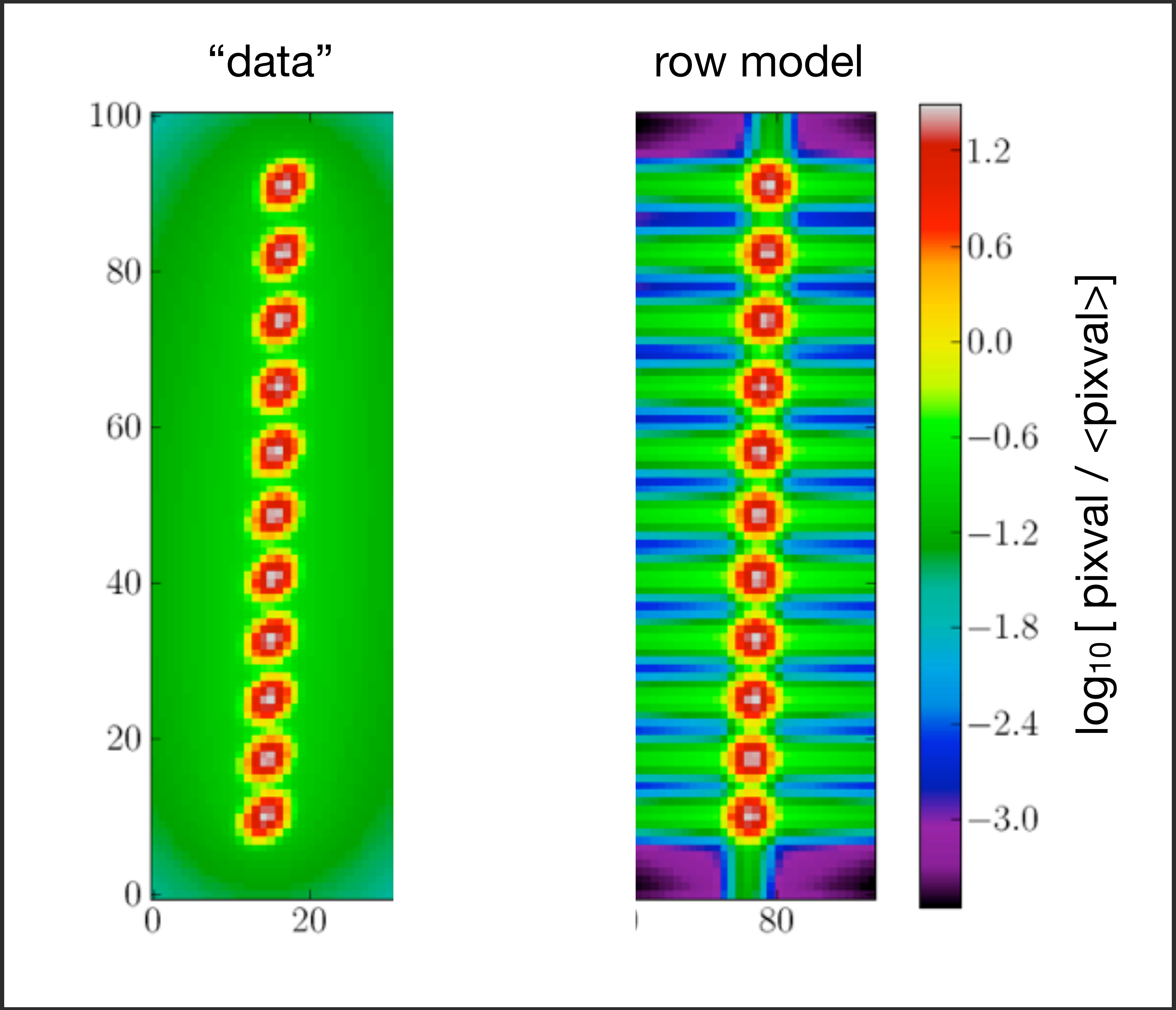
Row-by-row optimal extraction can only be correct when the spectrograph PSF is a *separable* function of x and y.

2D PSF extraction correct for arbitrary PSF shape.

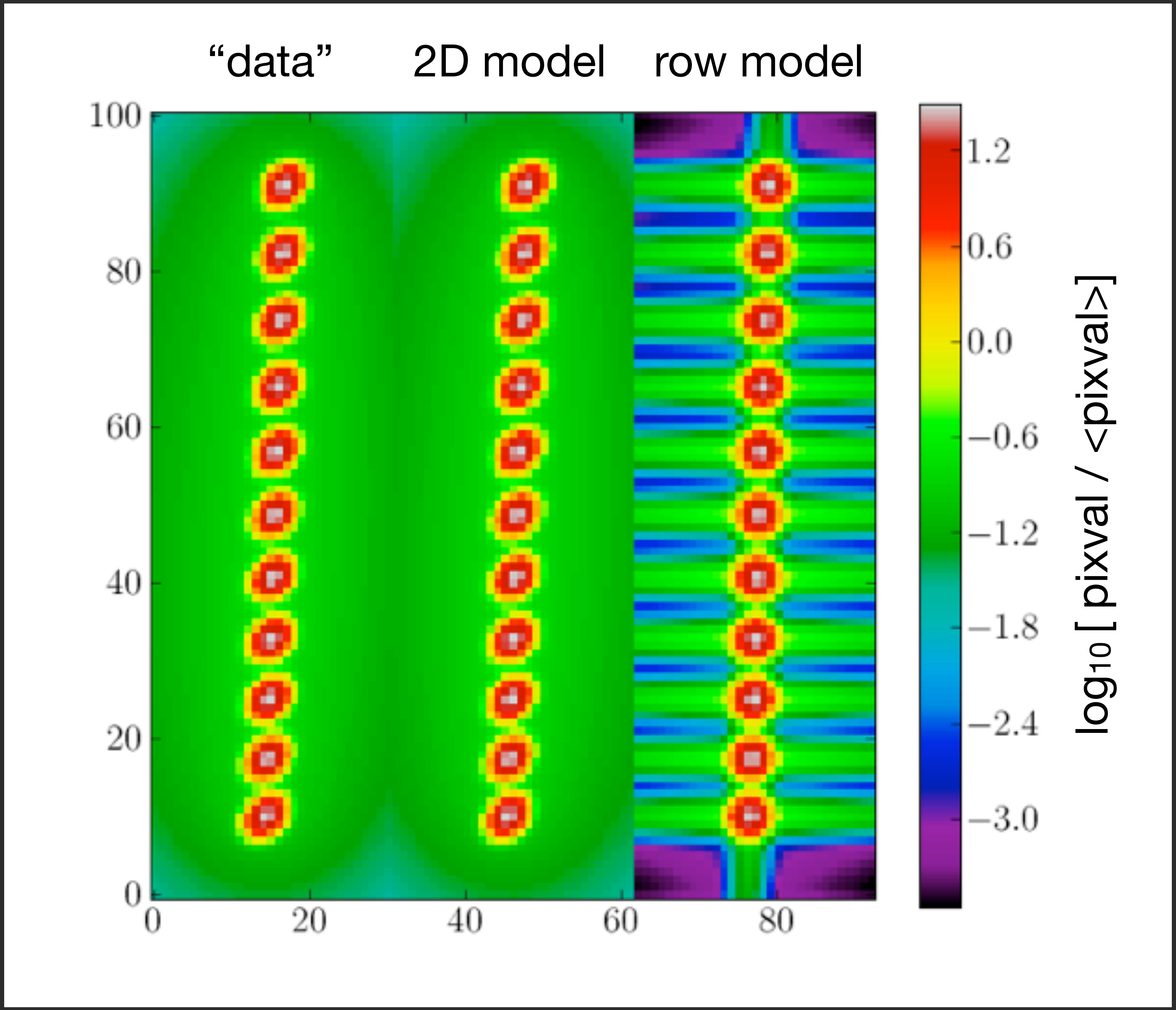
Extraction as image modeling



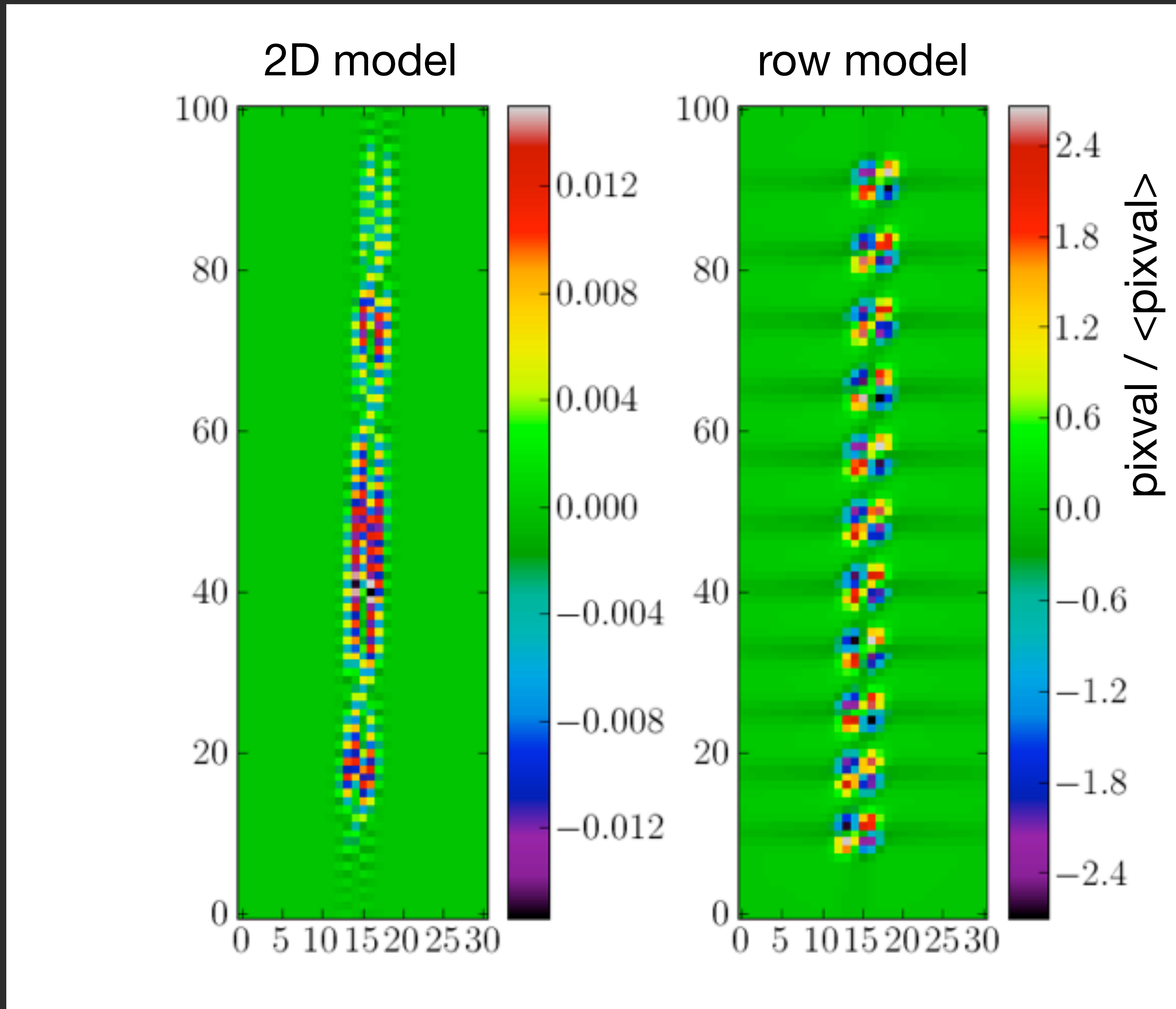
Extraction as image modeling



Extraction as image modeling



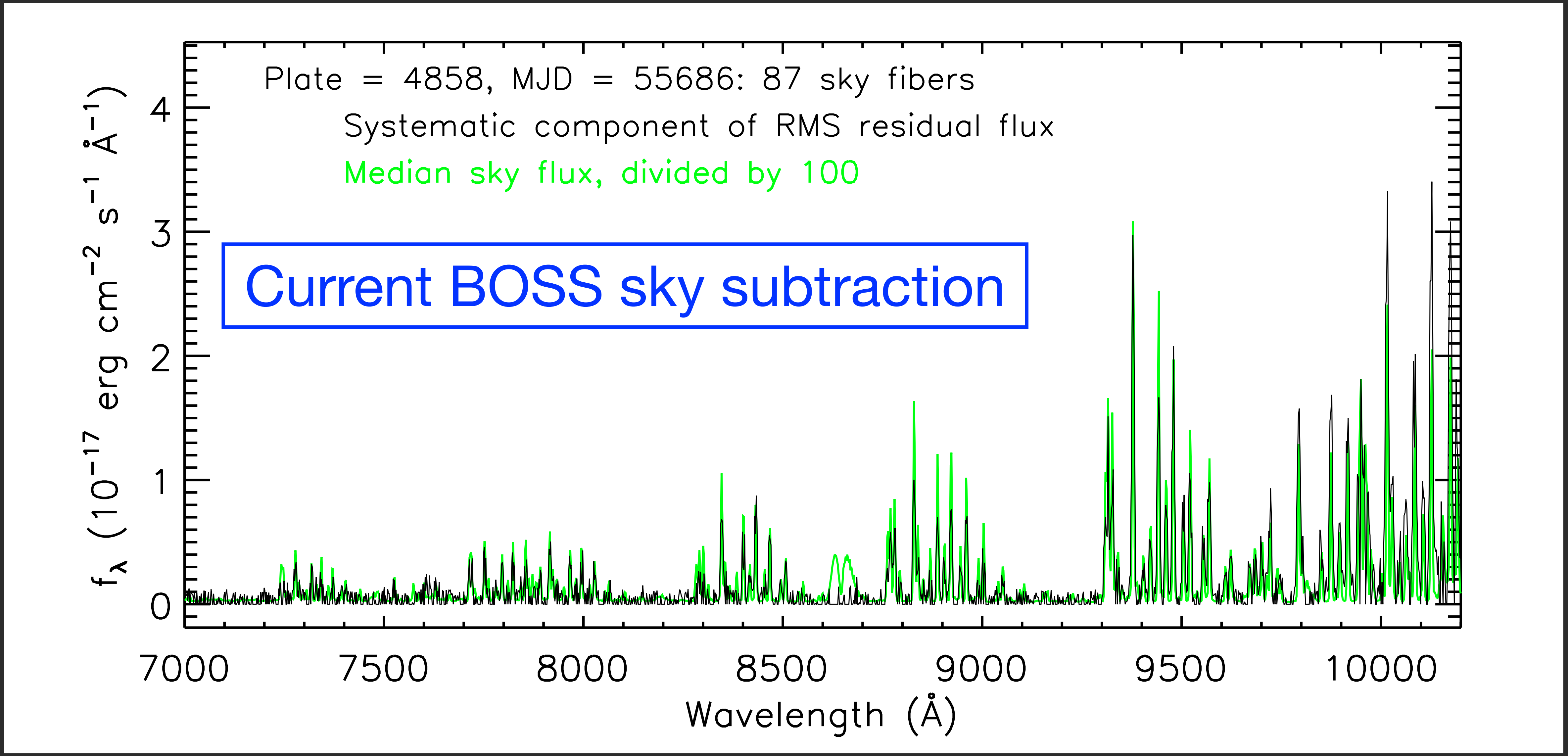
2D extraction model residuals



Why does this matter?

1) Poisson-limited sky subtraction

=> Current and future faint-galaxy redshift surveys
(E.g., BOSS, BigBOSS -- esp. [OII] ELG sample, ...)



Why does this matter?

1) Poisson-limited sky subtraction

=> Current and future faint-galaxy redshift surveys
(E.g., BOSS, BigBOSS -- esp. [OII] ELG sample, ...)

2) Extraction as lossless compression

=> All high-precision spectroscopic science
(Up to and including, e.g., RV planet surveys?)

What is a spectrum, anyway?

Not just

f = extracted spectrum vector

What is a spectrum, anyway?

Not just

  $\boldsymbol{f}$ = extracted spectrum vector

but also

  $\boldsymbol{R}$ = band-diagonal line-spread function matrix

and

  $\boldsymbol{C}$ = spectrum covariance matrix

What is a spectrum, anyway?

Not just

$\boldsymbol{f}$ = extracted spectrum vector

but also

$\boldsymbol{R}$ = band-diagonal line-spread function matrix

and

$\boldsymbol{C}$ = spectrum covariance matrix

Together, these encode the likelihood of a given input spectrum model ***m*** via:

$$\chi^2 (\boldsymbol{m} \mid \text{data}) = (\boldsymbol{f} - \boldsymbol{R}\, \boldsymbol{m})^T \boldsymbol{C}^{-1} (\boldsymbol{f} - \boldsymbol{R}\, \boldsymbol{m})$$

How do we do this?

Projection of input spectrum to CCD pixel frame of raw data via "calibration matrix" A

(CCD pixel counts) = ***A*** (input spectrum counts) + (noise)

(That is, A_{jk} = predicted counts in pixel "j" from monochromatic input at wavelength "k".

How do we do this?

Projection of input spectrum to CCD pixel frame of raw data via "calibration matrix" A

(CCD pixel counts) = ***A*** (input spectrum counts) + (noise)

(That is, A_{jk} = predicted counts in pixel "j" from monochromatic input at wavelength "k".

Generalizes and incorporates:

- Trace solution
- Wavelength solution
- 2D spectrograph PSF and its variation (i.e., aberrations)
- Relative and absolute throughput variation
- CCD pixel sensitivity variations
- Etc.

How do we do this?

Projection of input spectrum to CCD pixel frame of raw data via "calibration matrix" A

(CCD pixel counts) = ***A*** (input spectrum counts) + (noise)

(That is, A_{jk} = predicted counts in pixel "j" from monochromatic input at wavelength "k".

Likelihood of any model spectrum ***m*** then encoded by:

$$\chi^2 (\boldsymbol{m} \mid \boldsymbol{p}) = (\boldsymbol{p} - \boldsymbol{A}\,\boldsymbol{m})^T \boldsymbol{N}^{-1} (\boldsymbol{p} - \boldsymbol{A}\,\boldsymbol{m})$$

This is forward-modeling to the raw pixels.

How do we do this?

"De-convolved" minimum-χ^2 spectrum solution would be

$$\boldsymbol{m} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}\,(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1})\,\boldsymbol{p}$$

How do we do this?

"De-convolved" minimum-χ^2 spectrum solution would be

$$\boldsymbol{m} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}\,(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1})\,\boldsymbol{p}$$

Now define resolution R and covariance C via:

$$(\boldsymbol{A}^{\mathrm{T}}\,\boldsymbol{N}^{-1}\,\boldsymbol{A}) = \boxed{\boldsymbol{Q}\,\boldsymbol{Q}} = (\boldsymbol{R}^{\mathrm{T}}\,\boxed{\boldsymbol{C}^{-1}}\,\boldsymbol{R})$$

diagonal

Symmetric matrix root

Department of Physics and Astronomy
ASTRONOMY AND ASTROPHYSICS
THE UNIVERSITY OF UTAH

Adam S. Bolton

How do we do this?

"De-convolved" minimum-χ^2 spectrum solution would be

$$\boldsymbol{m} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}\,(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1})\,\boldsymbol{p}$$

Now define resolution R and covariance C via:

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A}) = \boxed{\boldsymbol{Q}\,\boldsymbol{Q}} = (\boldsymbol{R}^{\mathrm{T}}\boxed{\boldsymbol{C}^{-1}}\boldsymbol{R})$$

diagonal

Symmetric matrix root

And define extracted spectrum as:

$$\boldsymbol{f} = \boldsymbol{R}\,(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}\,(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1})\,\boldsymbol{p}$$

(Like a "re-convolution" of the de-convolved solution)

How do we do this?

Likelihood of any model spectrum $\boldsymbol{m}$ encoded by

$$\chi^2 (\boldsymbol{m} \mid \boldsymbol{f}) = (\boldsymbol{f} - \boldsymbol{R}\, \boldsymbol{m})^T \boldsymbol{C}^{-1} (\boldsymbol{f} - \boldsymbol{R}\, \boldsymbol{m})$$

is then mathematically equivalent to

$$\chi^2 (\boldsymbol{m} \mid \boldsymbol{p}) = (\boldsymbol{p} - \boldsymbol{A}\, \boldsymbol{m})^T \boldsymbol{N}^{-1} (\boldsymbol{p} - \boldsymbol{A}\, \boldsymbol{m})$$

(up to a constant offset)

How do we do this?

Likelihood of any model spectrum ***m*** encoded by

$$\chi^2 (\boldsymbol{m} \mid \boldsymbol{f}) = (\boldsymbol{f} - \boldsymbol{R}\, \boldsymbol{m})^T \boldsymbol{C}^{-1} (\boldsymbol{f} - \boldsymbol{R}\, \boldsymbol{m})$$

is then mathematically equivalent to

$$\chi^2 (\boldsymbol{m} \mid \boldsymbol{p}) = (\boldsymbol{p} - \boldsymbol{A}\, \boldsymbol{m})^T \boldsymbol{N}^{-1} (\boldsymbol{p} - \boldsymbol{A}\, \boldsymbol{m})$$

(up to a constant offset)

Forward-modeling to a spectrum extracted in this manner is information-equivalent to forward-modeling to the raw CCD pixels.

What is extraction?

Calibration:
 Likelihood functional determination

Extraction:
 Likelihood functional compression

Measurement:
 Likelihood functional projection

Summary of 2D PSF extraction

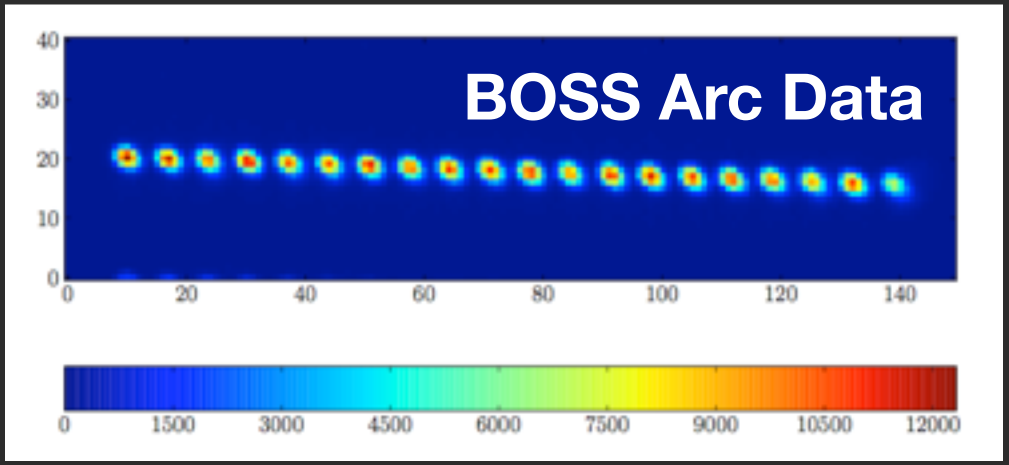
Major Advantages:

- Extraction as lossless compression
- Mathematically correct even for non-separable PSF
- Incorporates explicit model of 2D data
- Poisson-limited sky subtraction
- Data products "look & feel like spectra"

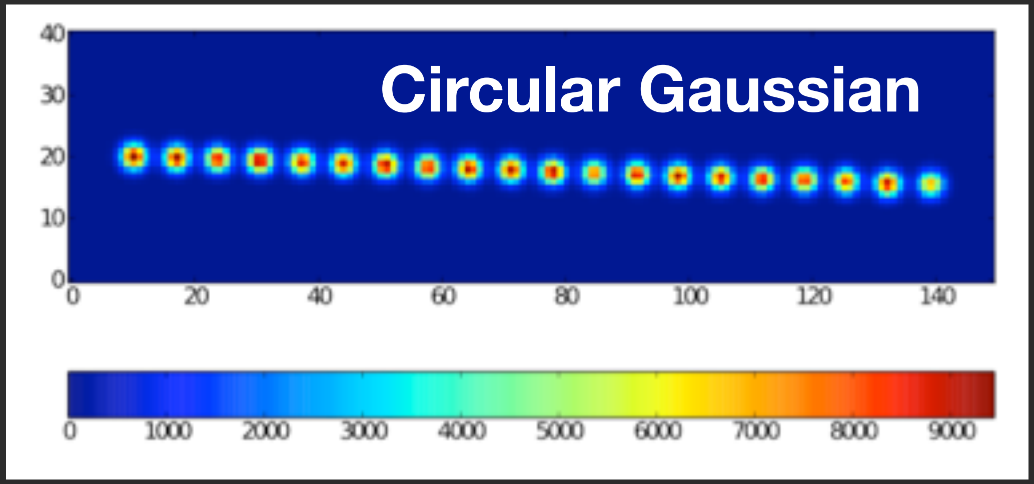
Major Challenges:

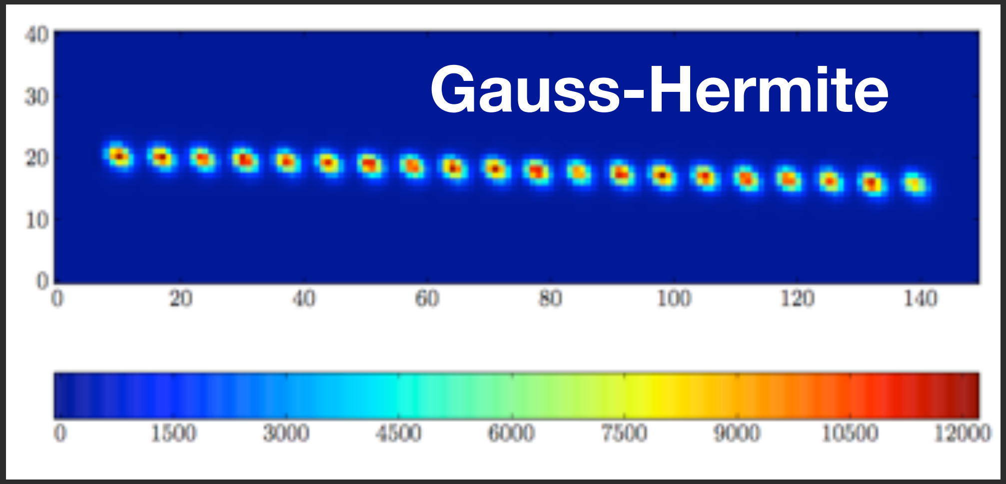
- Extraction coupled across wavelengths
- Requires exquisite calibration
- Some subtlety related to flux normalization

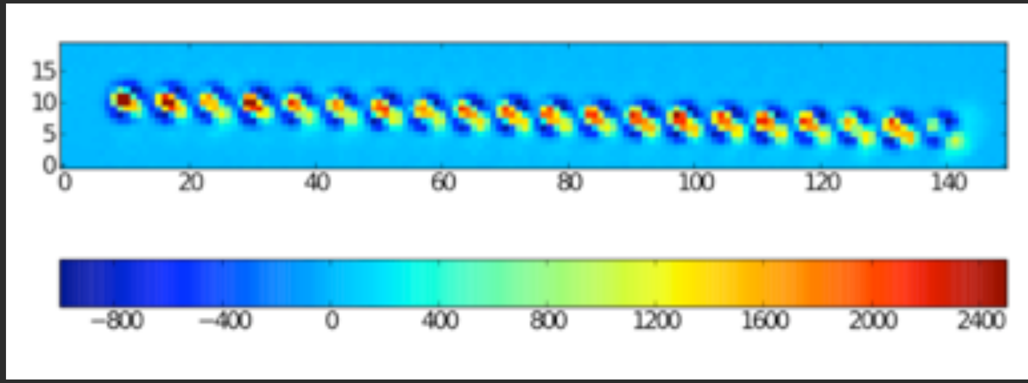
Development & Implementation Status

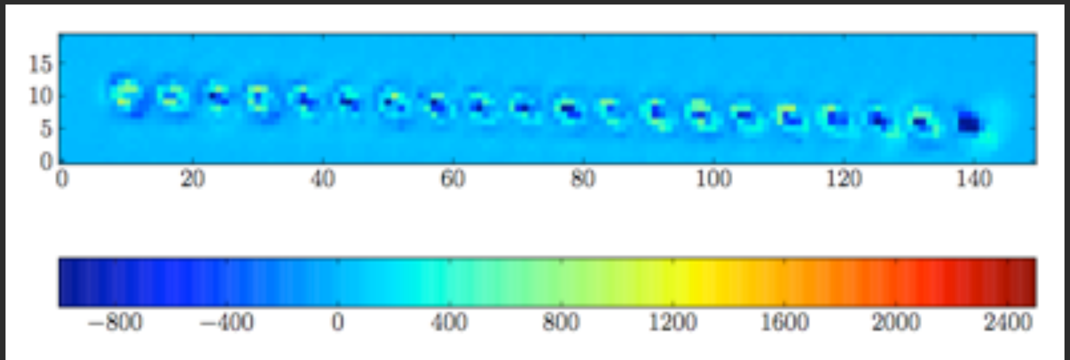


Images from
Parul Pandey
M.S. Thesis
U. of Utah







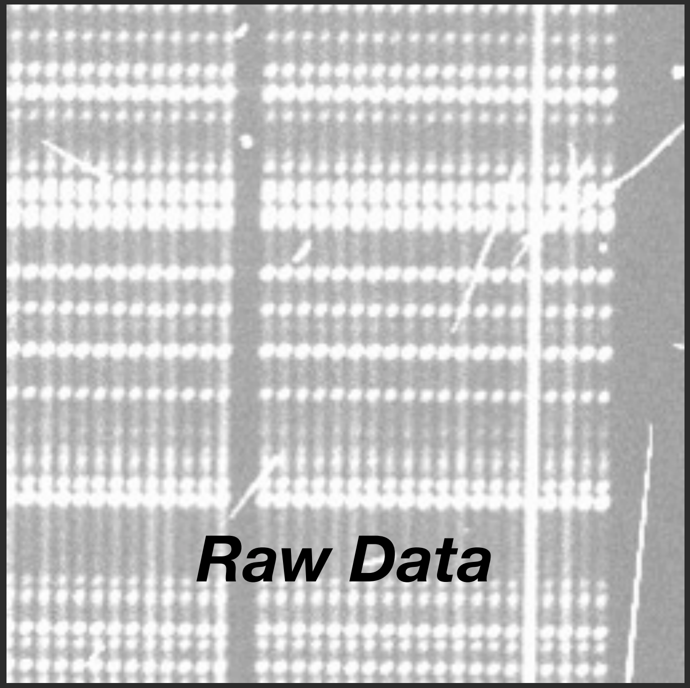


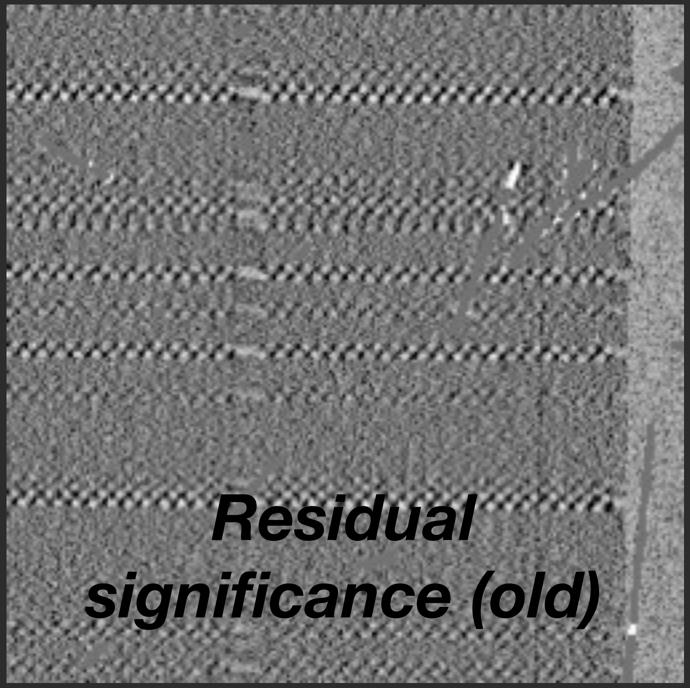
Also: wing component, higher order GH, pixelized PSF

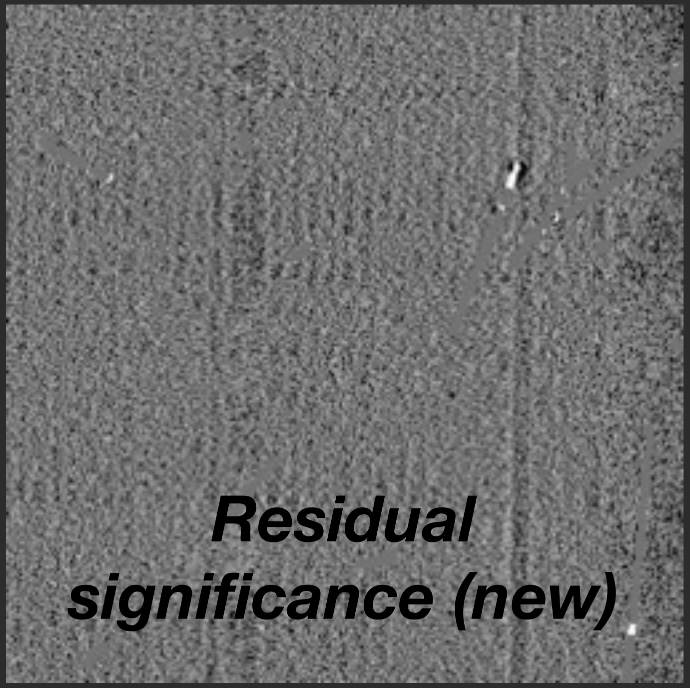
Development & Implementation Status

Demonstrated path for computational tractability:

- Decompose among bundles, exposures, spectrographs, and wavelength ranges







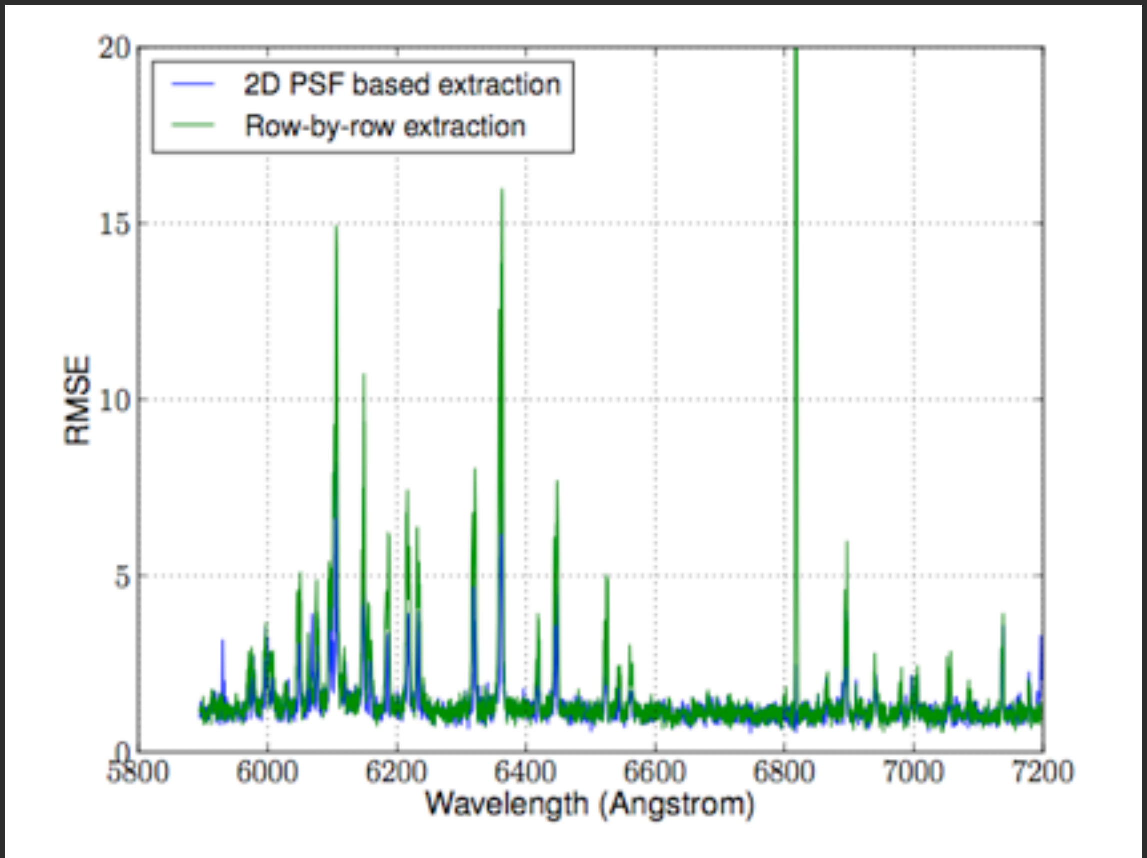
Effort in summer 2011 and ongoing by:

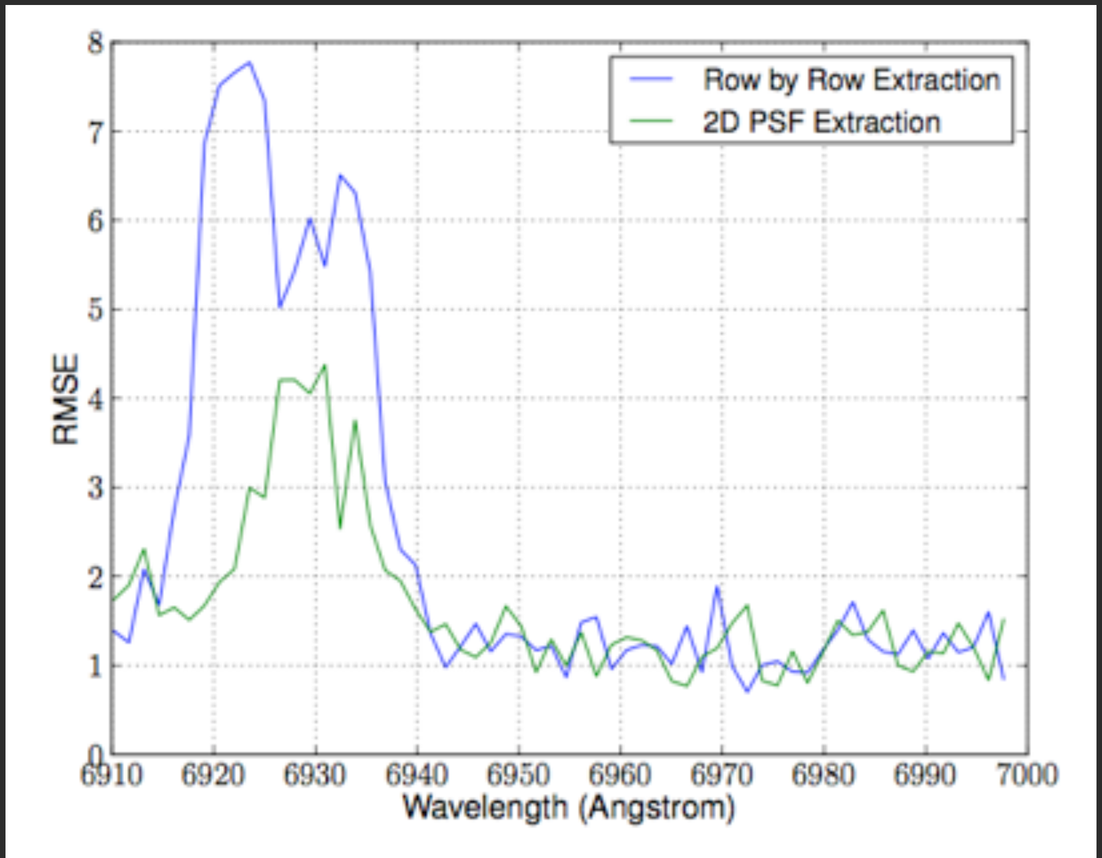
ASB, Joel Brownstein, Parul Pandey (U. of Utah)

Stephen Bailey, Ted Kisner, David Schlegel (LBNL)

Benefits in extracted-spectrum frame

Sky subtraction, as simulated by arc-lamp data





Images from Parul Pandey, M.S. Thesis 2011, U. of Utah

Software Requirements on Hardware

Separability: we absolutely need gaps between bundles of fibers where cross-talk goes to zero

True resolution: metric is not camera spot EE or flux-weighted r^2, but *wavelength autocorrelation of PSF:*

$$[\int p(x,y;\lambda)\, p(x,y;\lambda+\Delta\lambda)\, dx\, dy] / [\int p^2(x,y;\lambda)\, dx\, dy]$$

(N.B.: Rayleigh criterion is autocorrelation of 1/4)

Calibration: tunable monochromatic system for mapping out system calibration matrix?

Stability: fractional spectrum bias for assuming wrong PSF q(x,y) instead of right PSF p(x,y) is:

$$b = 1 - [\int p(x,y)\, q(x,y)\, dx\, dy] / [\int p^2(x,y)\, dx\, dy]$$

Software Requirements on Hardware

Ultimately calls for a full integration of data analysis software with instrumental design software

=> Optimize *scientific* metrics in hardware design

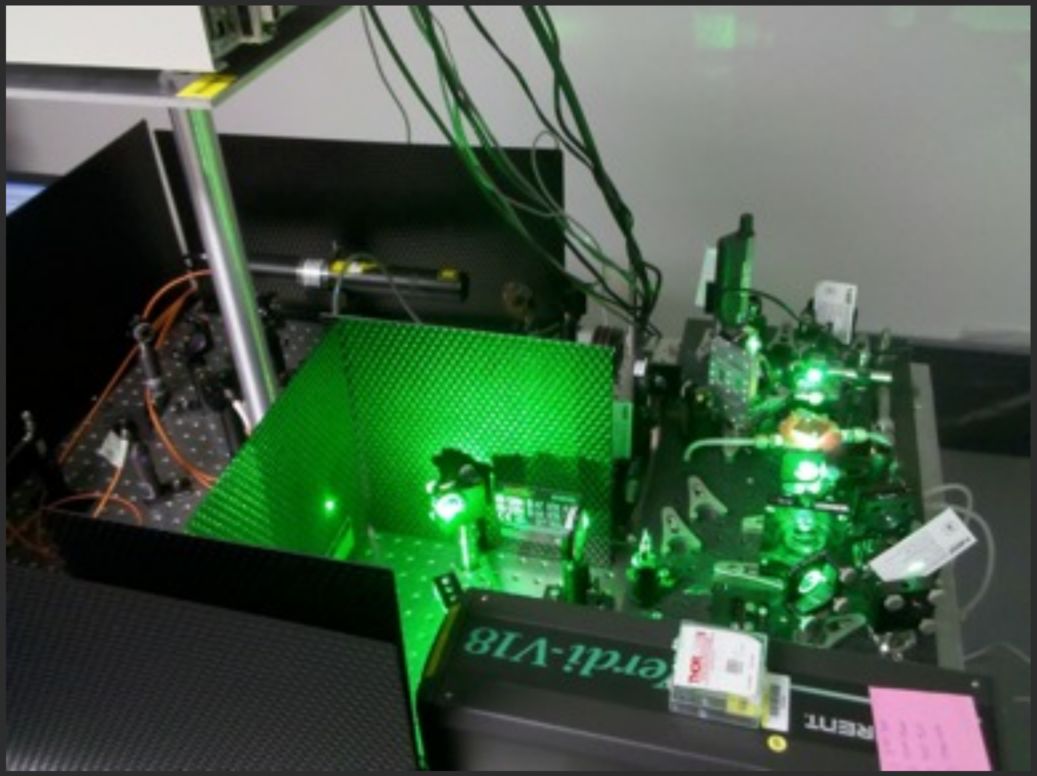
=> Tune instrument directly from science CCD data

=> “Use what you know” during analysis

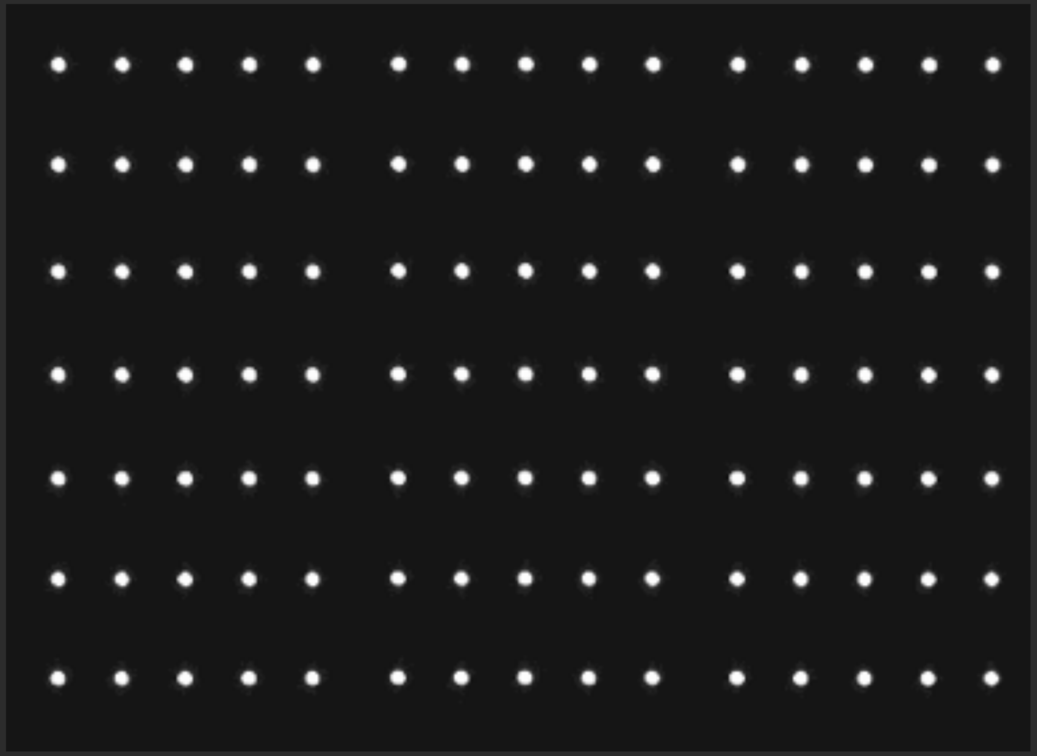
Monochromatic calibration

NIST-BOSS tunable laser experiment (w/ C. Cramer, K. Lykke)

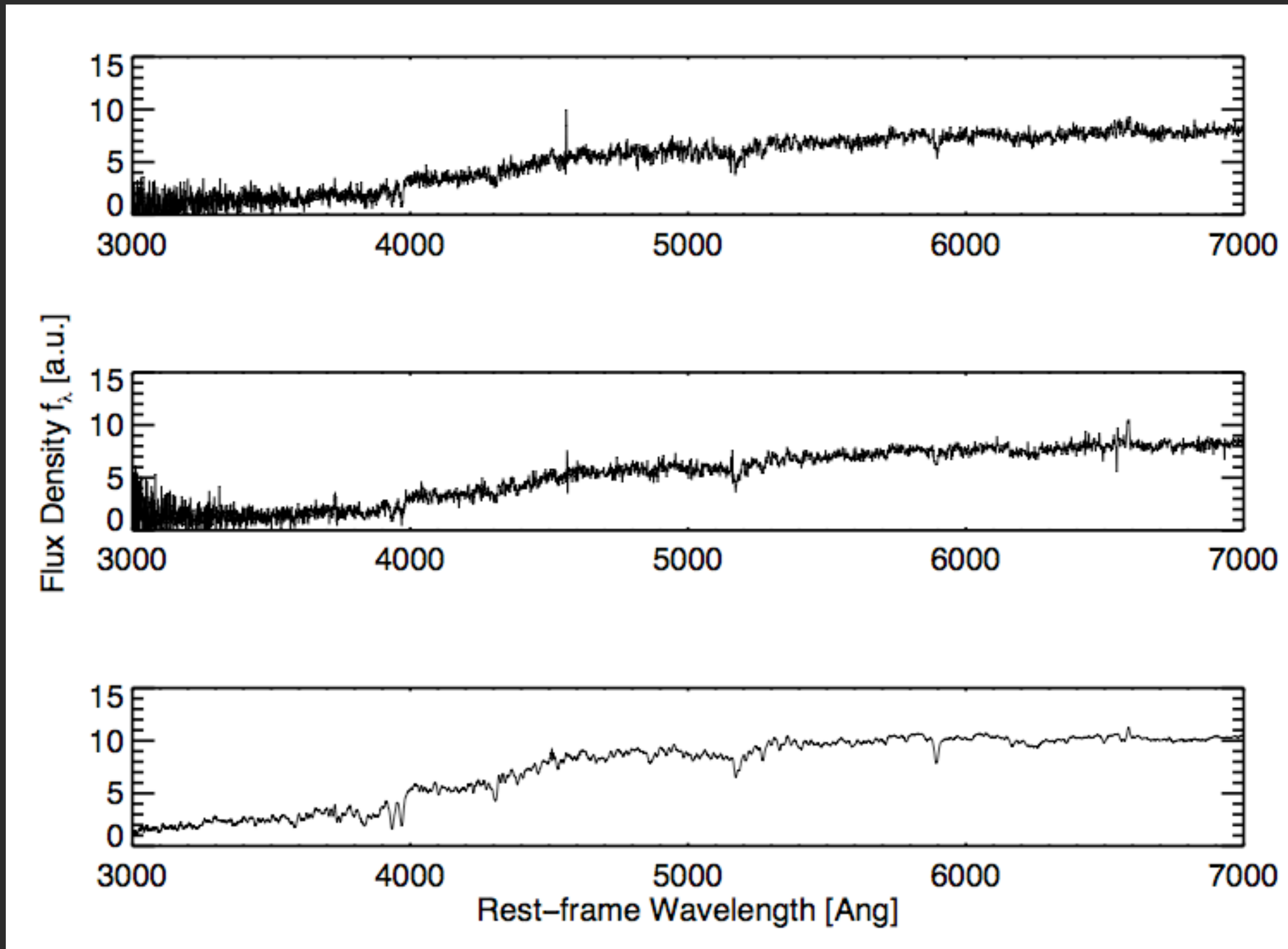
(Also see G. Tarle "Line-O-Matic")



vs.



Application: Bayesian stacking



Shu, ASB, et al., submitted (arXiv 1109.6678)

Application: Bayesian stacking

Model vdisp distribution at fixed z and M as a log-normal distribution (c.f. Bernardi et al. 2003):

$$p(log_{10}\sigma|m,s) = \frac{1}{\sqrt{2\pi}s} e^{-\frac{(log_{10}\sigma - m)^2}{2s^2}} \qquad (1)$$

Constrain parameters in (z, M) bins by integrating over all spectra and all vdisp values:

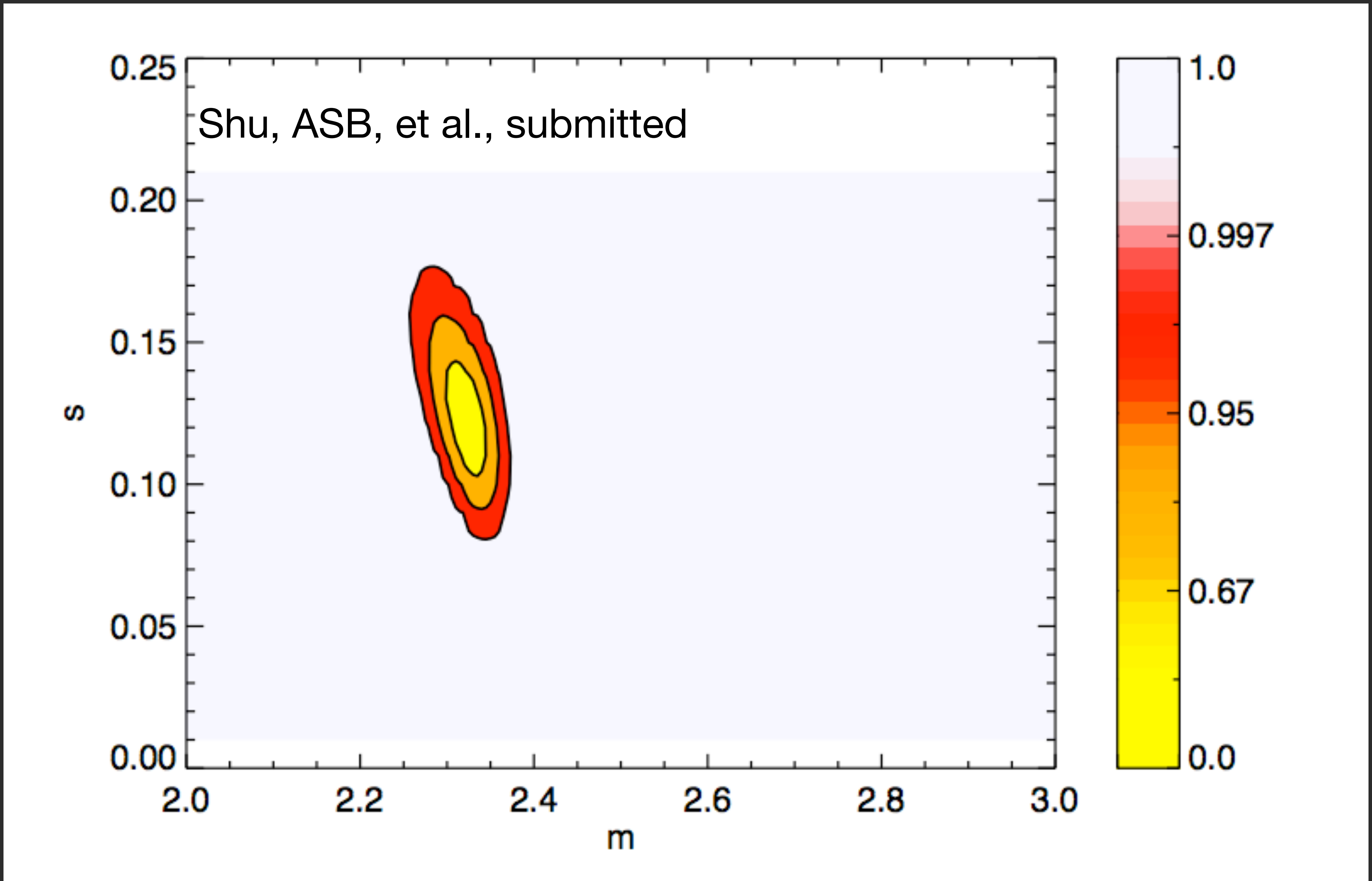
$$\begin{aligned}\mathcal{L}(m,s|\{\vec{d}\}) &= p(\{\vec{d}\}|m,s) \\ &= \prod_i p(d_i|m,s) \qquad (2) \\ &= \prod_i \int \mathrm{d}log_{10}\sigma \, p(d_i|log_{10}\sigma) \, p(log_{10}\sigma|m,s)\end{aligned}$$

$$p(m,s|\{\vec{d}\}) = \frac{p(\{\vec{d}\}|m,s)p(m,s)}{p(\{\vec{d}\})} \qquad (3)$$

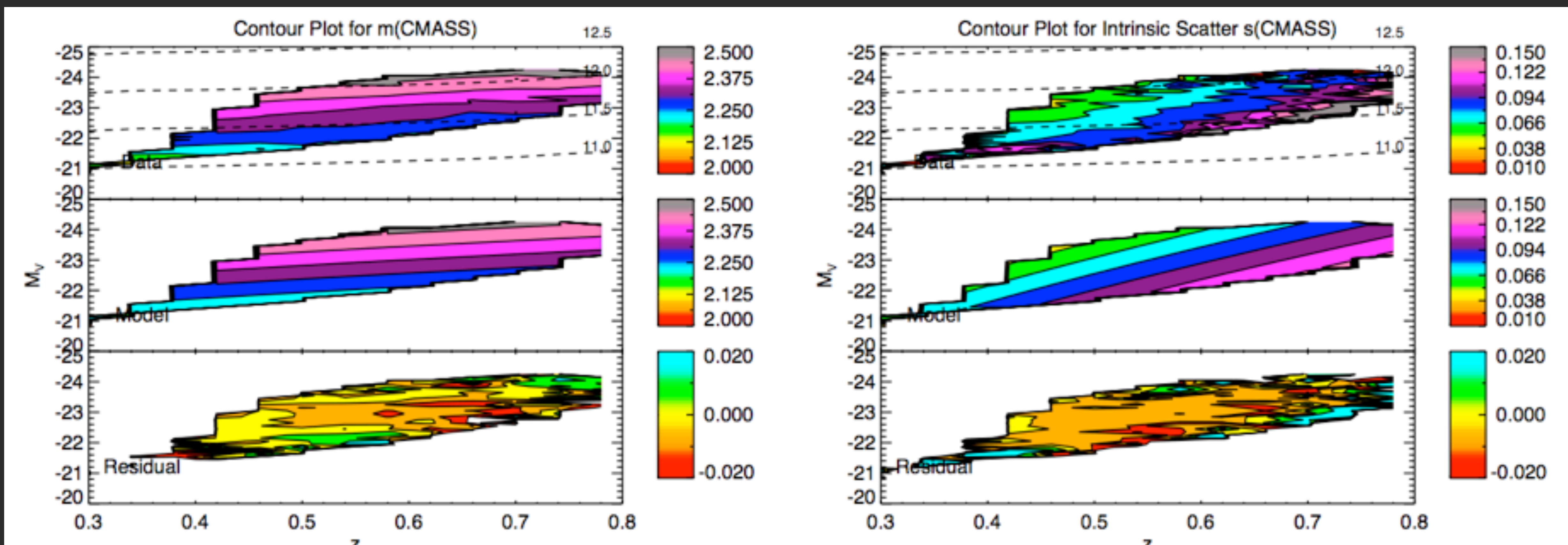
Shu, ASB, et al., submitted (arXiv 1109.6678)

N.B.: if you stack directly, you will measure $\sigma = 10$^[m + s^2 ln(10)]

Posterior probability for a single bin



Distribution results: population evolution



Shu, ASB, et al., submitted

Summary and conclusions

- Full 2D forward modeling of raw data is the way of the future for spectroscopic extraction
- Poisson-limited sky subtraction for ground-based faint-galaxy redshift surveys (BOSS, BigBOSS)
- Lossless compression of spectrum likelihood functional
- We have the algorithmic framework, and are currently putting it into practice
- Major challenges are in calibration, computation, and integration of data analysis with hardware design

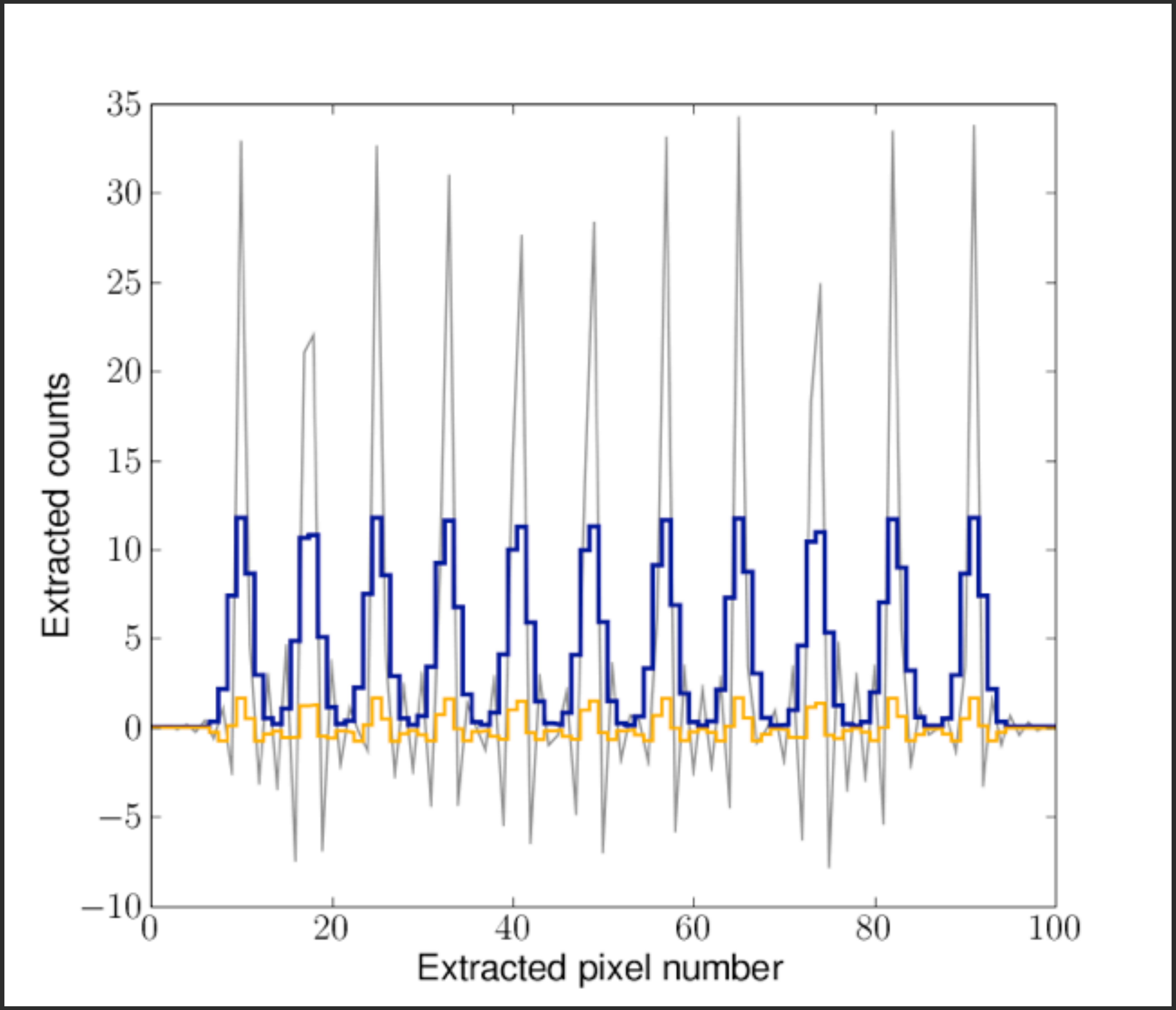
Thank You!

Department of Physics and Astronomy
ASTRONOMY AND ASTROPHYSICS
THE UNIVERSITY OF UTAH

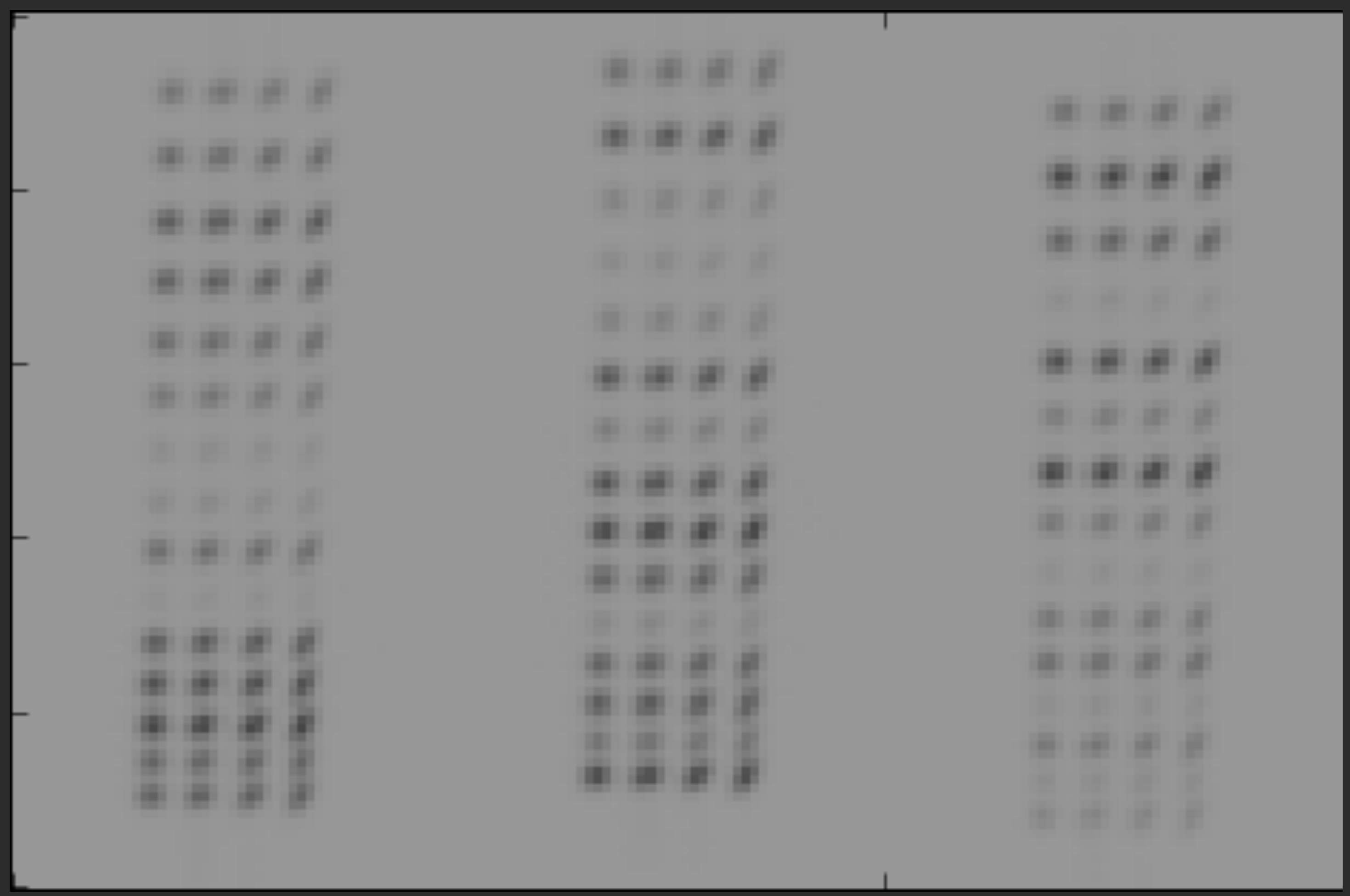
ADASS XXI 2011-11-09

Adam S. Bolton

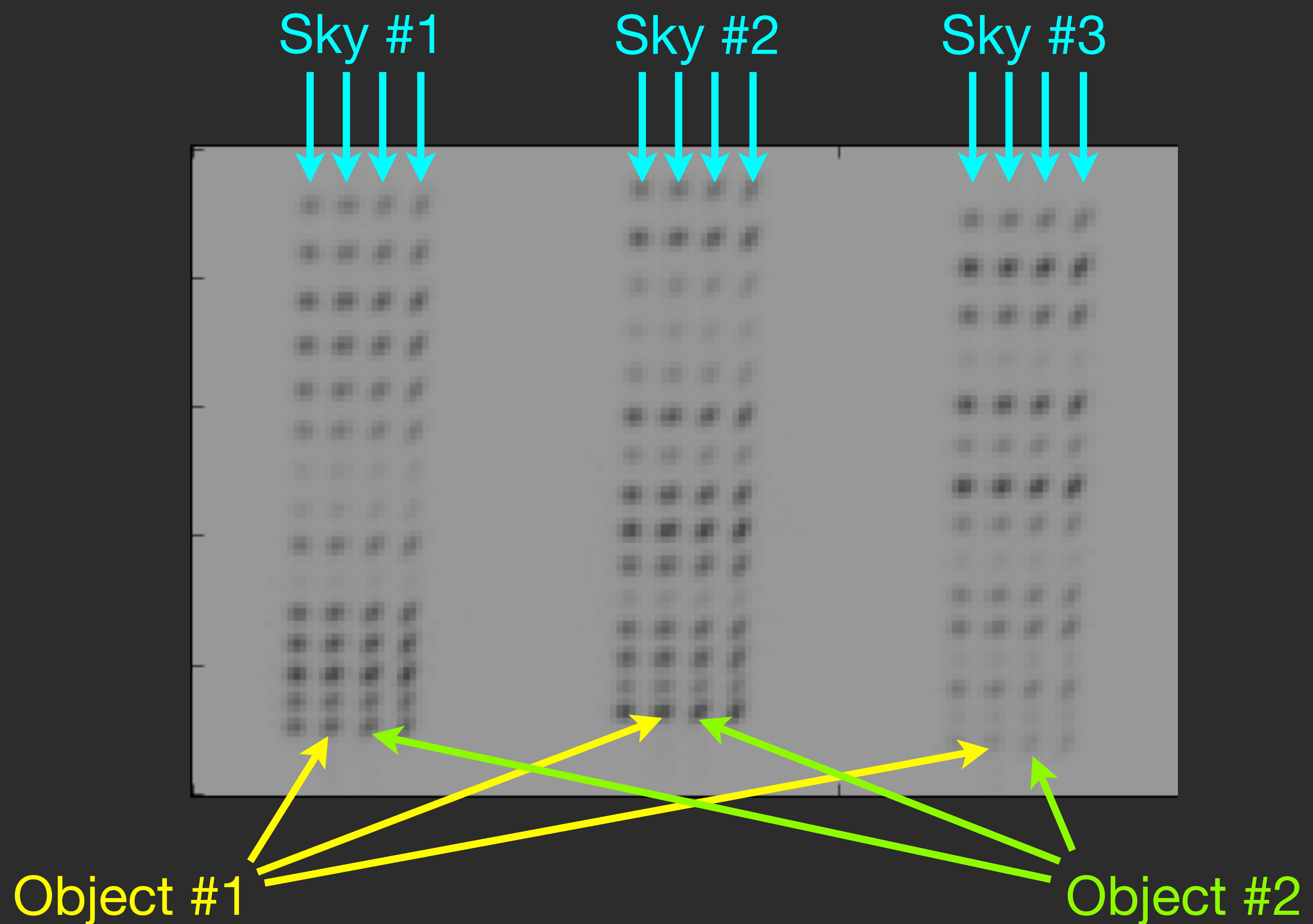
Deconvolution and reconvolution



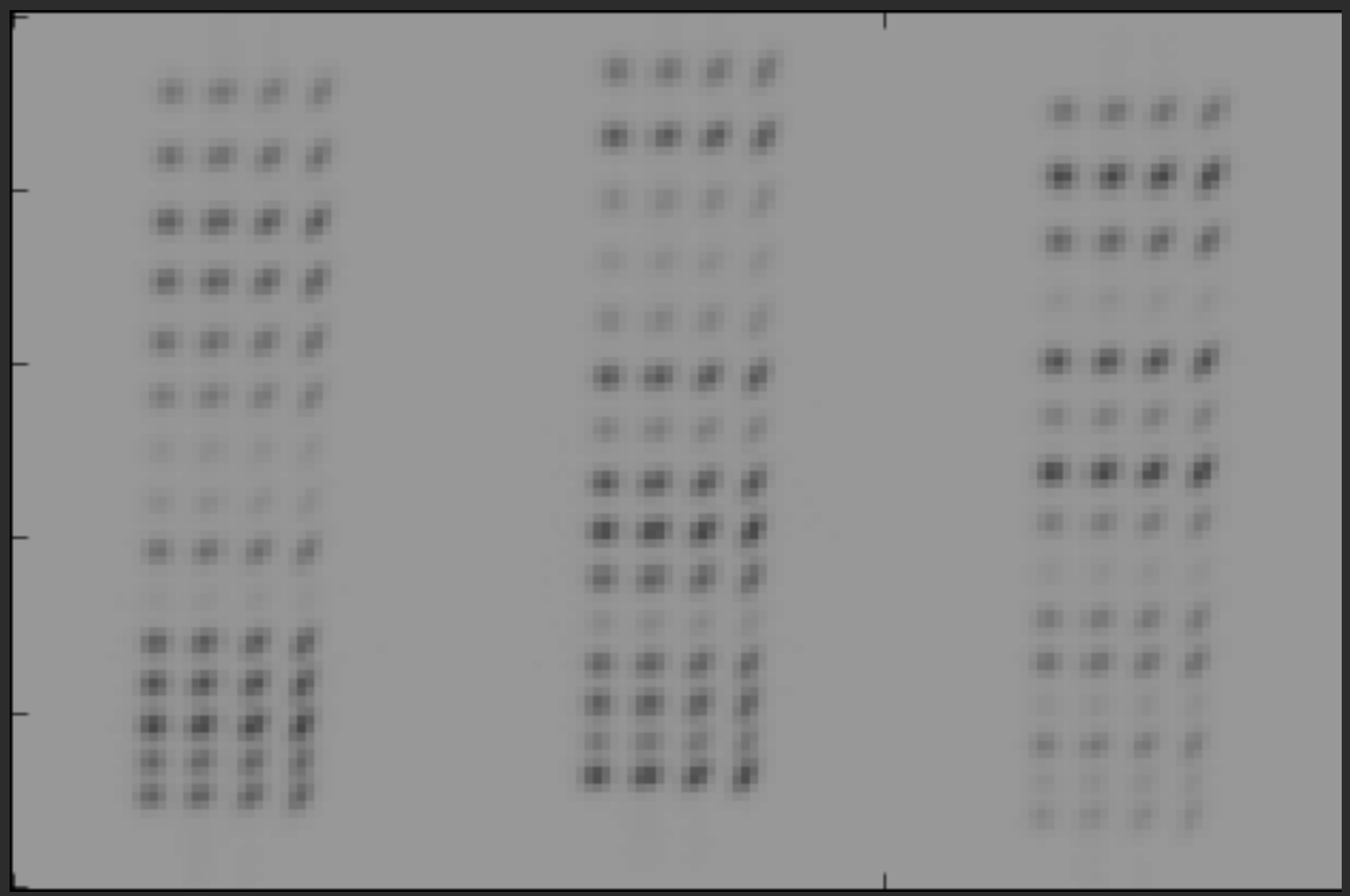
Multi-frame, multi-fiber simulated data



Multi-frame, multi-fiber simulated data



Multi-frame, multi-fiber simulated data

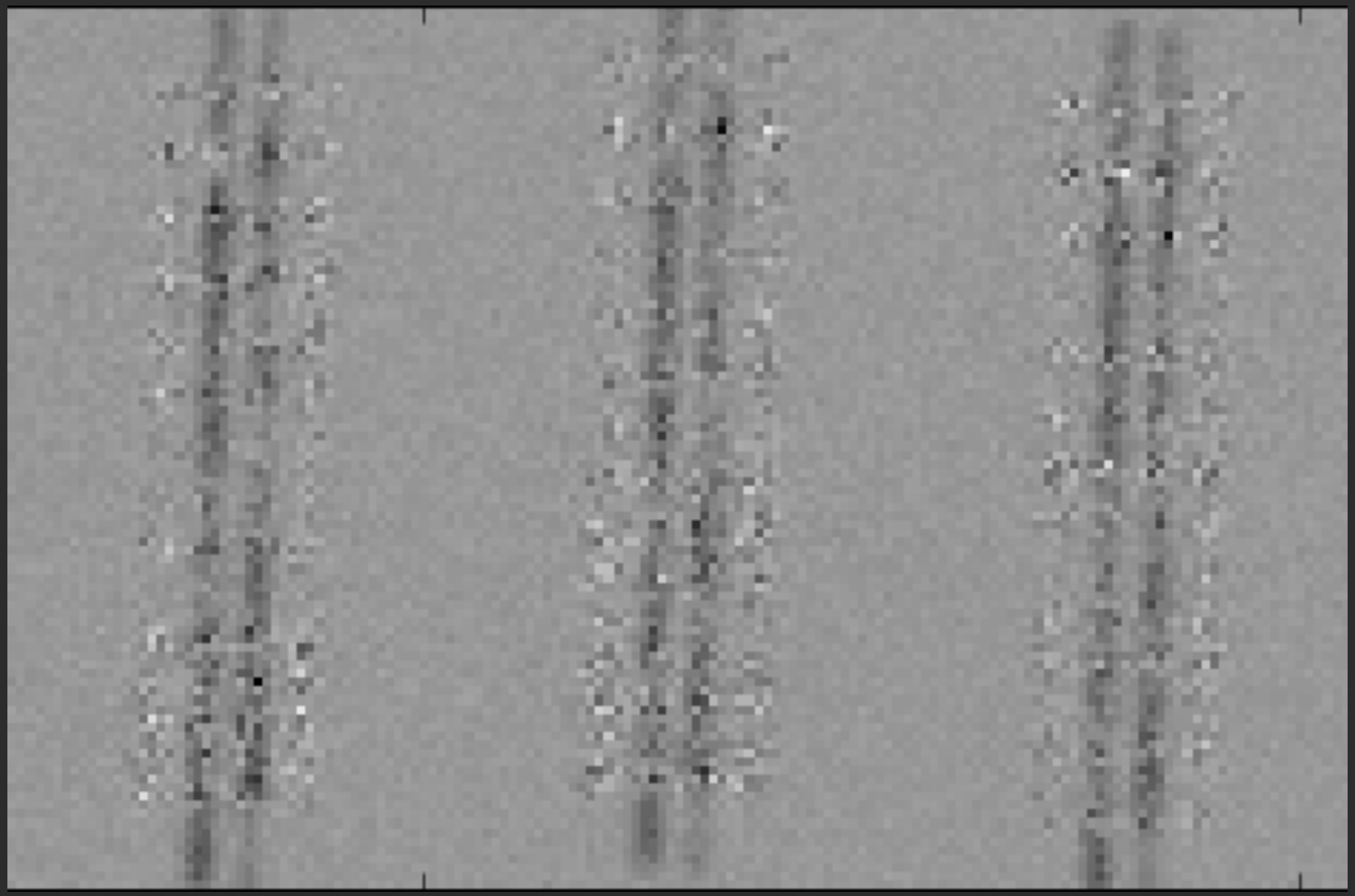


Objflux = Skyflux / 20

ObjSNR ≈ 5 (per extracted sample, sky-noise limited)

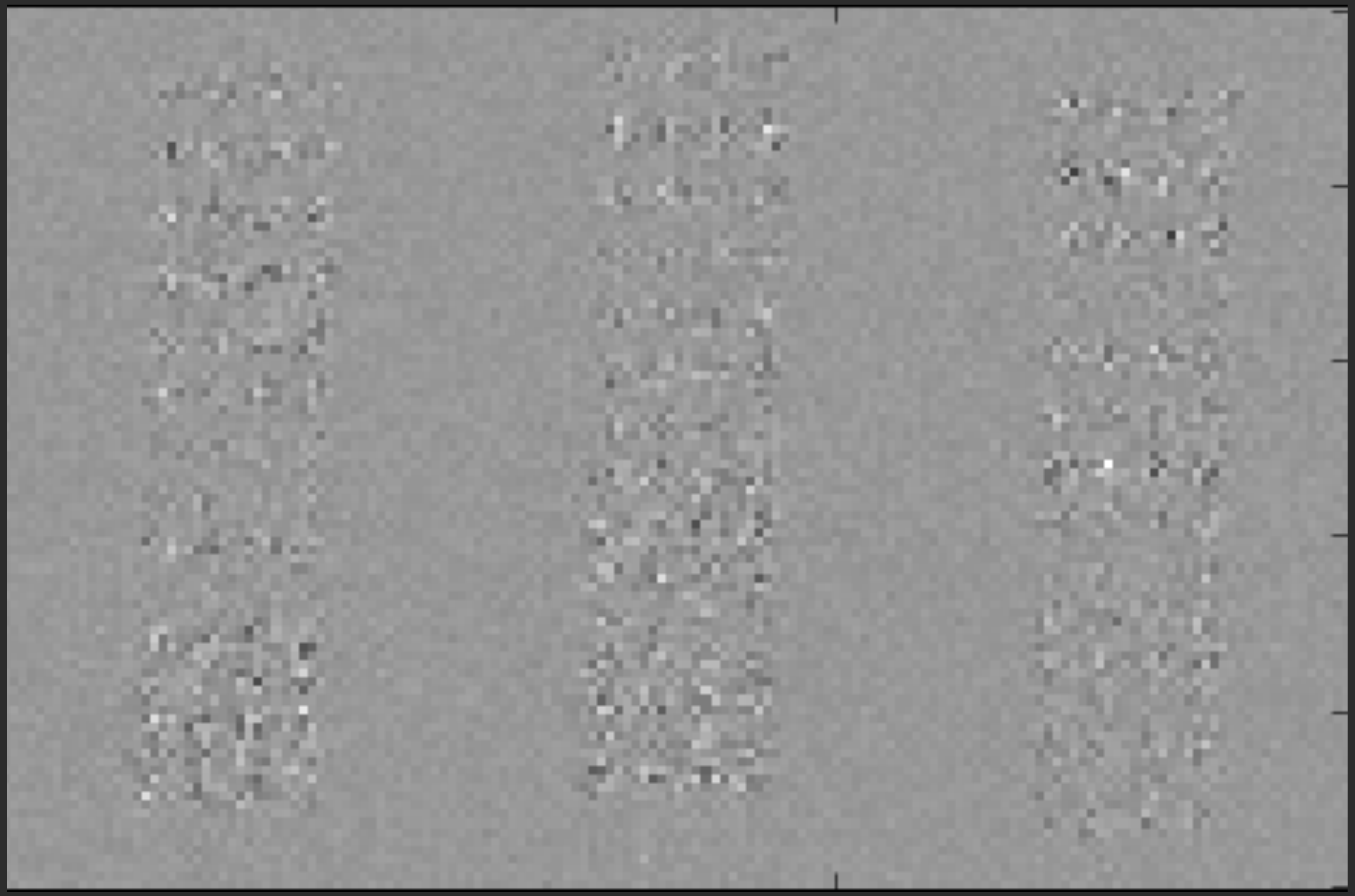
Sky model decomposed & removed

Sky spectrum is modeled "upstream" from optical heterogeneities between fibers



(Grayscale stretch X 40 relative to previous)

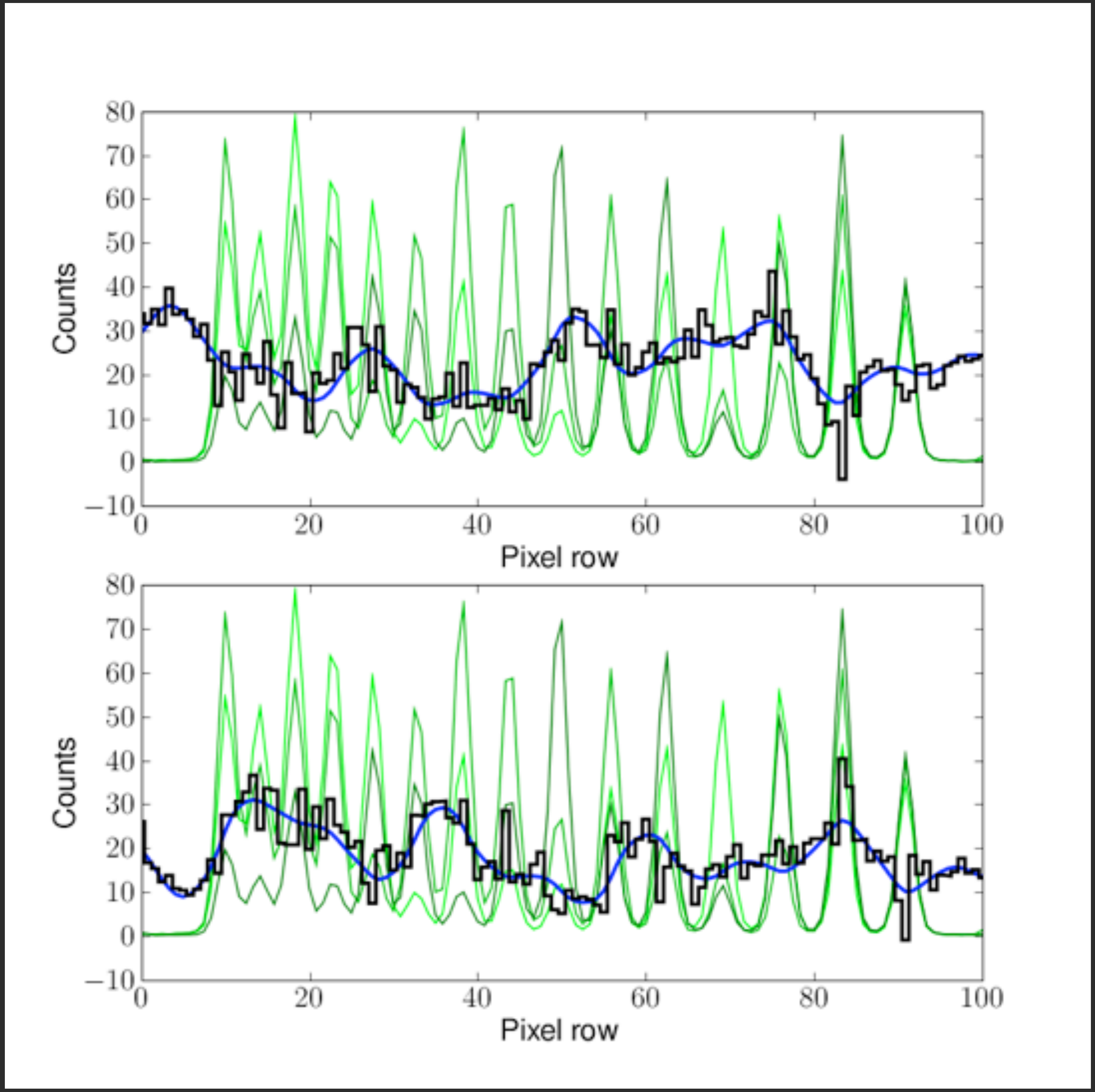
All models removed



Consistent with pure noise

Extracted objects + skies

Sky scaled down by a factor of 20 in plot



RMS error-scaled residuals of unity