SAND2015-10888C

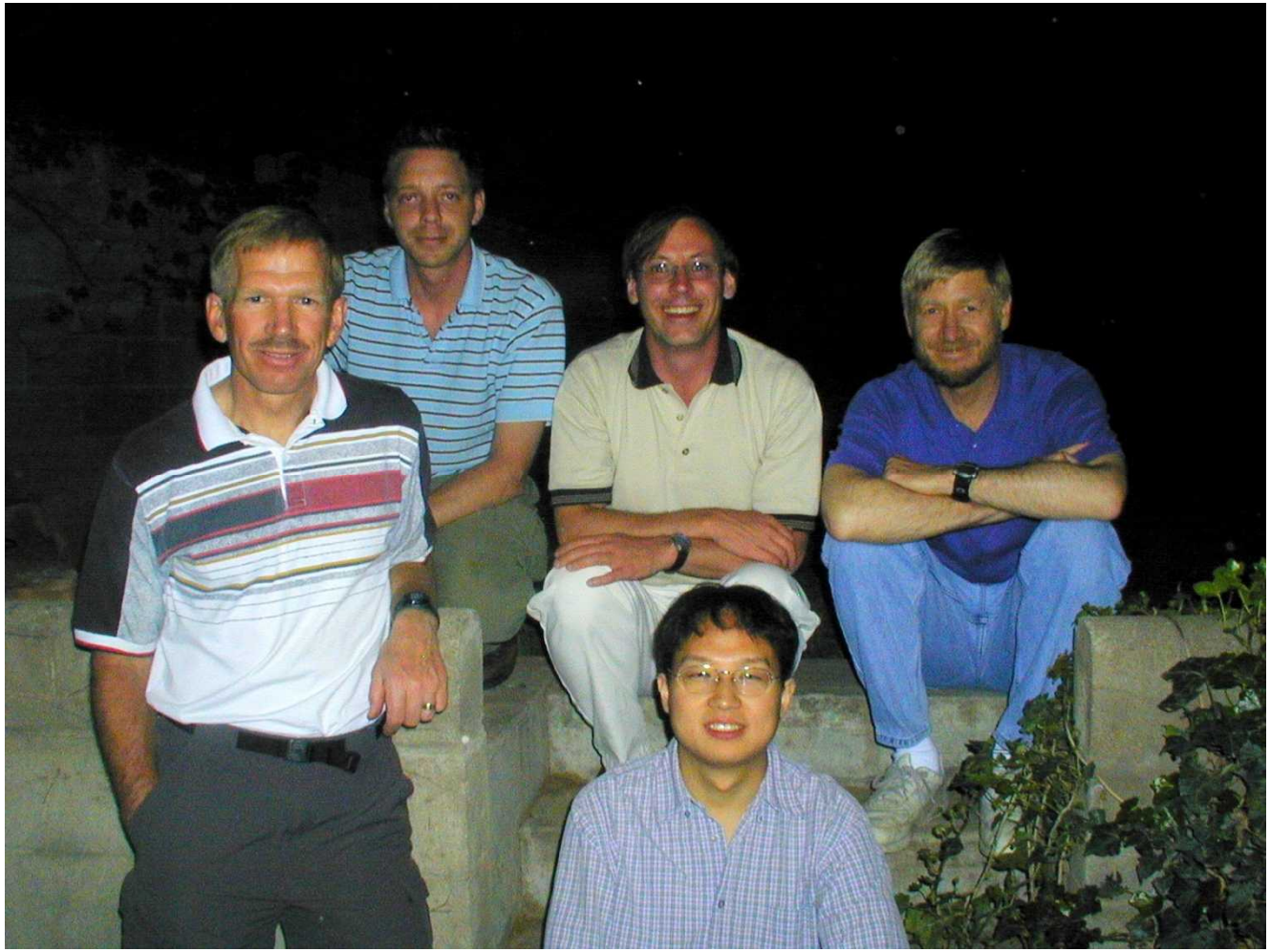
Automating Multivariate Curve Resolution for Success

David Haaland,[1] Howland Jones,[2] David Melgaard,[3] Jerilyn Timlin[3]

[1]Spectral Resolutions, Albuquerque, NM
[2]HyperImage Solutions, Rio Rancho, NM
[3]Sandia National Laboratories, Albuquerque, NM

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Peter's Sabbatical In NM (2003)



Sandia National Laboratories

Congratulations Peter!

Peter has made many significant contributions to the field of chemometrics. However, to quote Peter, “I have attached a reprint of one of my more significant contributions, even though it is now 10 years old.”

Sandia National Laboratories

Overview

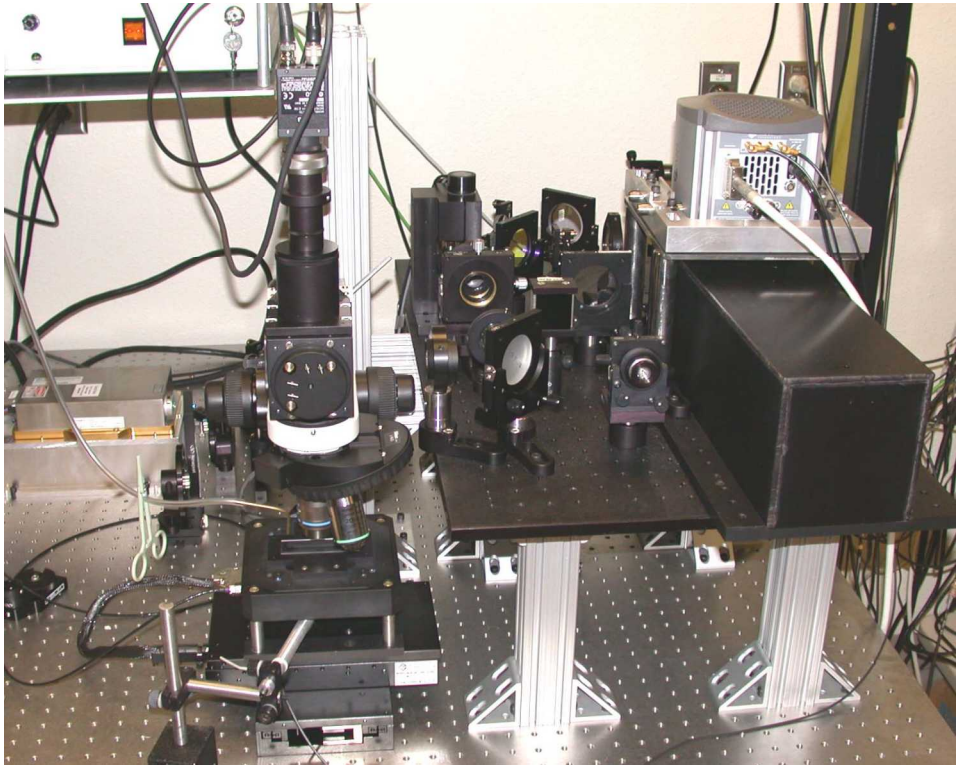
- 3D hyperspectral confocal fluorescence microscope
  - Brief description of microscope and multivariate curve resolution
- Discussion of approaches to automated MCR analyses*
  - Automated spectral preprocessing: despiking, offset and dark current removal
  - Automated ROI selection
  - Optional spatial compression and automatic PCA spectral compression
  - Automatic selection of optimal numbers of MCR pure components
  - Automated equality constraints on offset and dark spectral regions
- Demonstration of automated MCR analyses on simulated data
- Demonstration of automated MCR analyses on quantum dot images

*Jones *et al.* , "Preprocessing strategies to improve MCR analyses of hyperspectral images," Chemometrics and Intelligent Laboratory Systems (2012).

Sandia National Laboratories

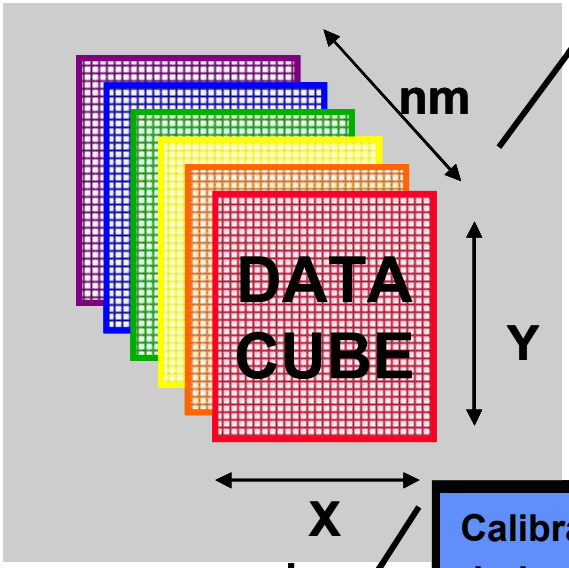
Sandia's Hyperspectral Confocal Fluorescence Microscope

- Fully confocal design
  - **High spatial resolution**
  - **Optical sectioning**
- High optical throughput
  - **Prism spectrometer**
  - **Electron multiplying CCD**
- Performance Specifications:
  - **488 nm laser excitation**
  - **10x, 20x, 60x, 100x objectives**
  - **Lateral Resolution = 0.25 μm**
  - **Axial Resolution = 0.60 μm**
  - **Spectral range 490-800 nm**
  - **Spectral resolution = 1-3 nm**
  - **Acquisition rate = 8300 spectra/sec**
  - **Rugate filter blocks laser and light from 1st 30 spectral channels**



Sinclair, Haaland, Timlin, and Jones
"Hyperspectral confocal microscope"
Applied Optics, **45**, 6283-6291 (2006).

Data Analysis Flowchart



512 wavelength channels
210 x 210 x 18 voxels
~800 Mb, ~1 min

WHAT?

Spectral Components

WHERE?
HOW MUCH?

Concentration Maps

Calibrate, despike, subtract dark, weight data for noise, PCA analysis

Preprocess

Random #s, literature, PCA results – normalized to unit length

Initial Estimates

MCR

Constrained alternating least squares analysis

Linear additive model applies
$D=CS^T$
# of comps known

Rigorous non-negativity & equality constraints
Efficient algorithms
Un-weight results

Sandia National Laboratories

Characteristics of Hyperspectral Images from EMCCD Detectors

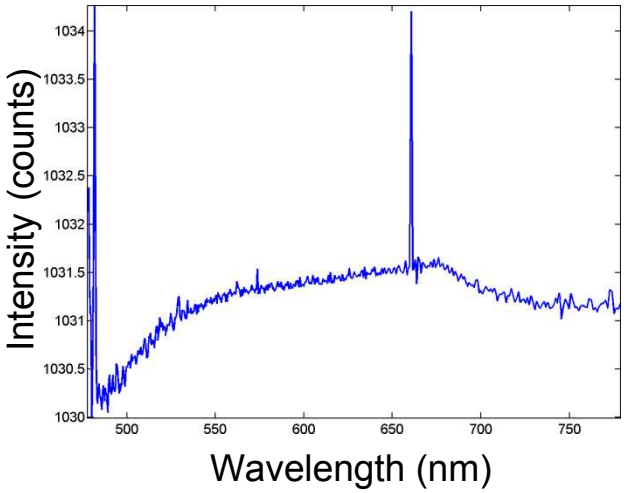
- Poisson noise is present
  - ($\sigma^2 \propto S$)
- Read noise is present
  - σ^2 varies with wavelength
  - High degree of spectral correlation
  - Randomly distributed spatially
- Spectral offset and structured spectral noise (dark current) is present
- Presence of all the above in the data complicates MCR analyses

Standard Practices in MCR Analysis

- Perform PCA on hyperspectral data
- Select number of independent spectral components based on the plot of log(EV) vs PCA factors (Scree plot)
- Truncate the PCA factors to be the same as the number of pure-component spectra selected from the Scree plot
- Perform MCR on PCA truncated image spectra assuming that the noise is iid(0,σ^2)

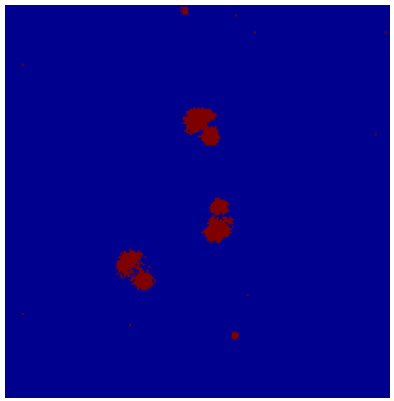
Preprocessing Steps Hyperspectral Image of Algae Cells

Raw Mean Spectrum



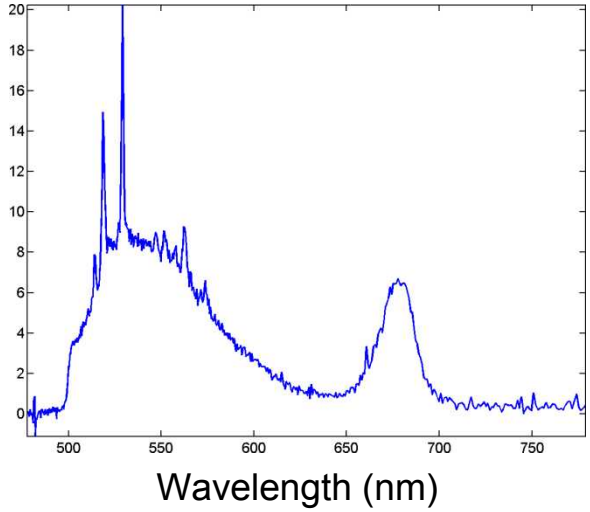
Despike, offset and dark current removal, spatial mask applied

Preprocessing



Pixel Selection

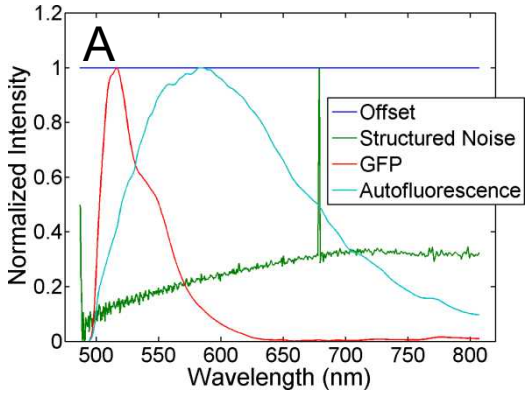
Preprocessed Mean Spectrum

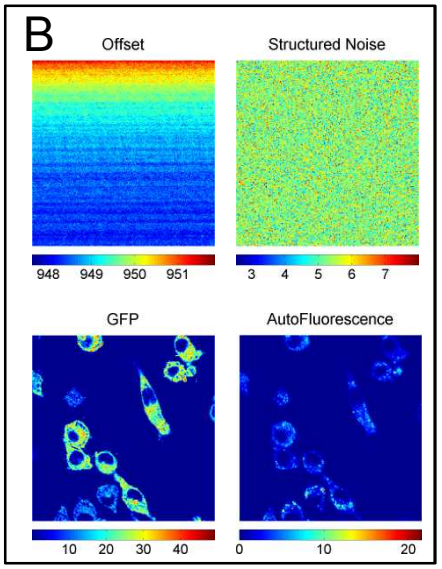


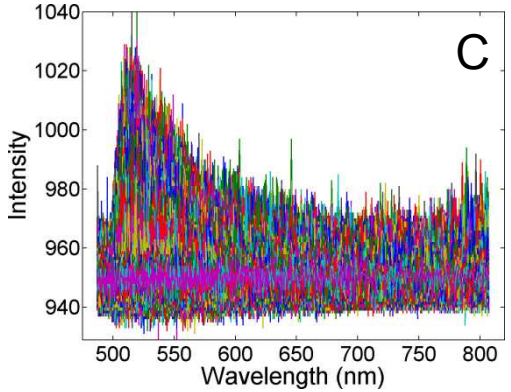
Note: All these preprocessing corrections can be made using 20 of the first 30 dark spectral channels.

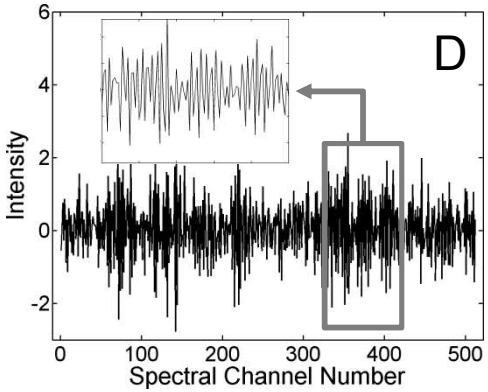
Sandia National Laboratories

Construction of Simulated Data RelA-GFP in Macrophage Cells



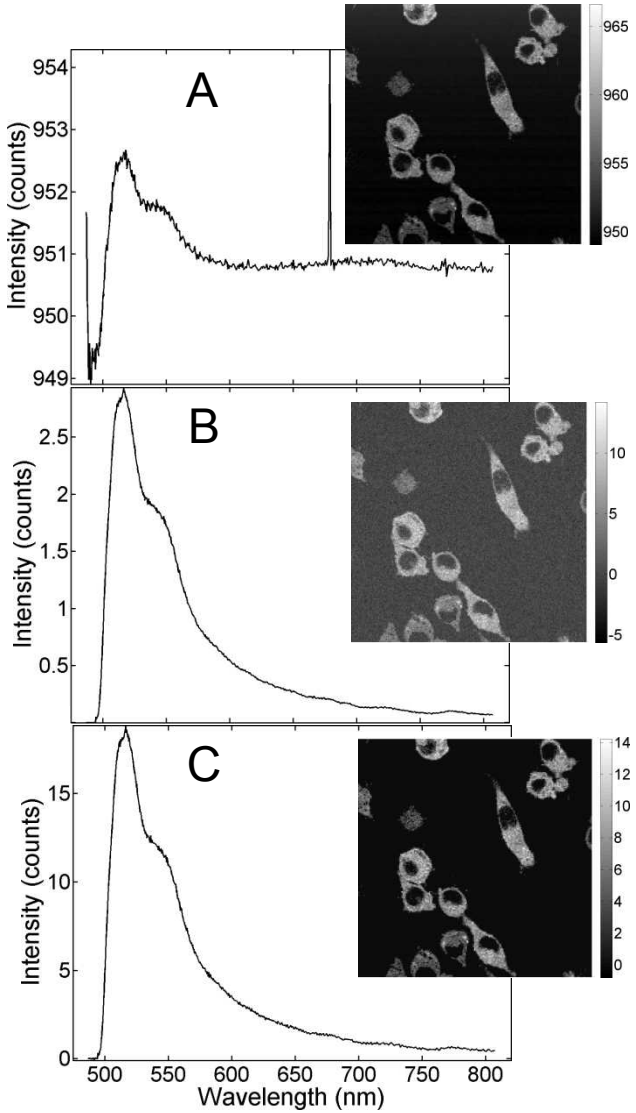






Sandia National Laboratories

Preprocessing of Simulated Hyperspectral Image Data



Macrophage cells with RelA GFP

A) No preprocessing

B) Stage 1 preprocessing: gamma ray spike removal, offset removal, dark removal

C) Stage 2 preprocessing: spatial pixel selection

First 20 dark spectral channels used for all preprocessing steps except despiking

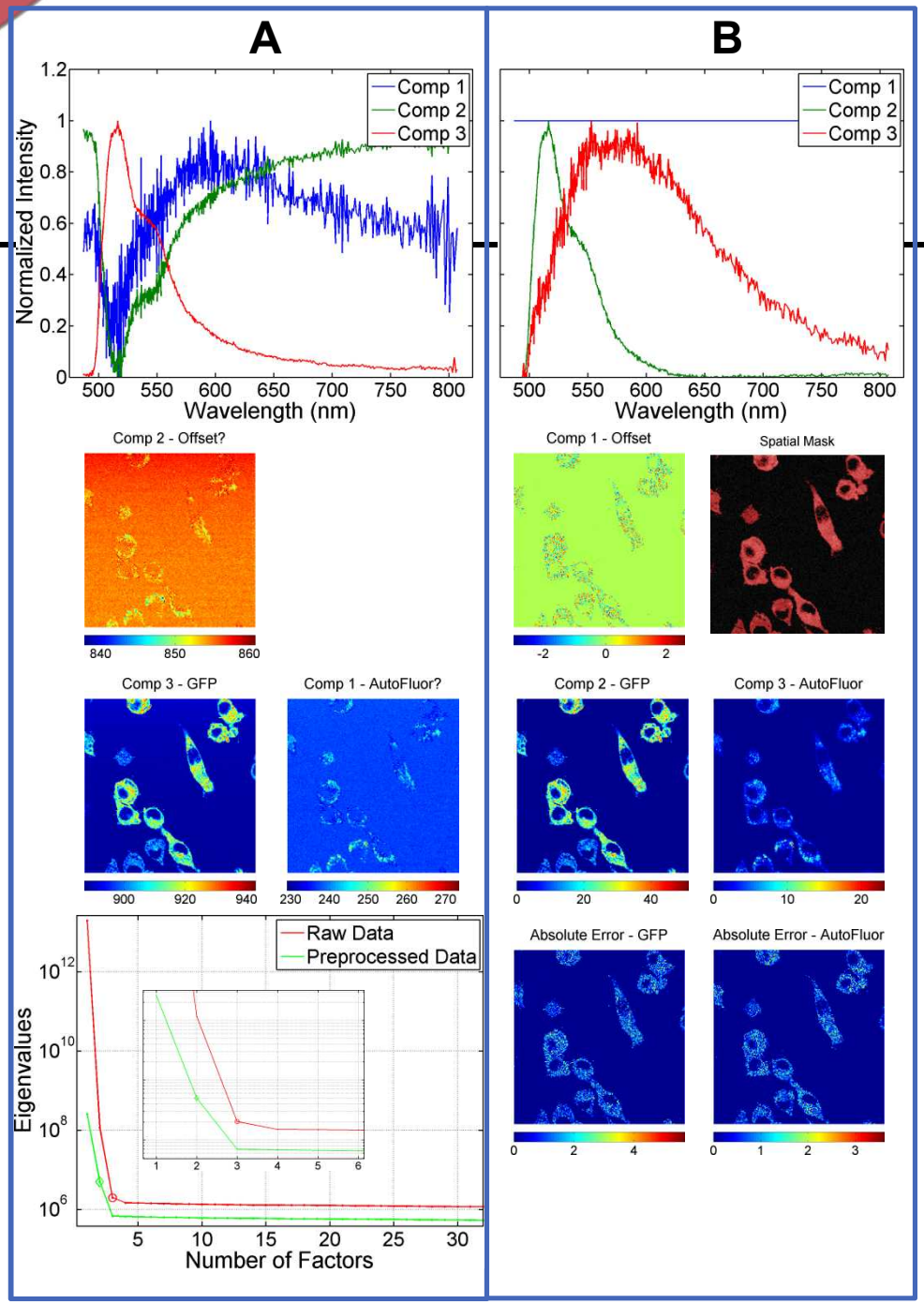
1) **Offset and dark removal: Fit first 20 dark channels to dark spectrum and offset to determine how much offset and dark current is in each spectrum**
2) **Determine spectral pixels with intensity 3σ above RMS of 20 dark channels**

Sandia National Laboratories

Post Preprocessing: Achieving Automated MCR Image Analysis

- Optional spatial compression of the image
  - Average adjacent spectral pixels (e.g., 4 x 4)
  - Minimizes correlated read noise
  - Improves spectral signal to noise
- PCA spectral compression
  - We retain 64 PCA factors
  - Automated selection of MCR components based upon statistical tests of eigenvalues
- Equality constraints
  - Zero the first 20 dark channels and equality constrain them to prevent mixing with the baseline and other components
  - Add offset and equality constrain it to fit residual offset (remove nonnegative concentration constraint)

MCR Analyses



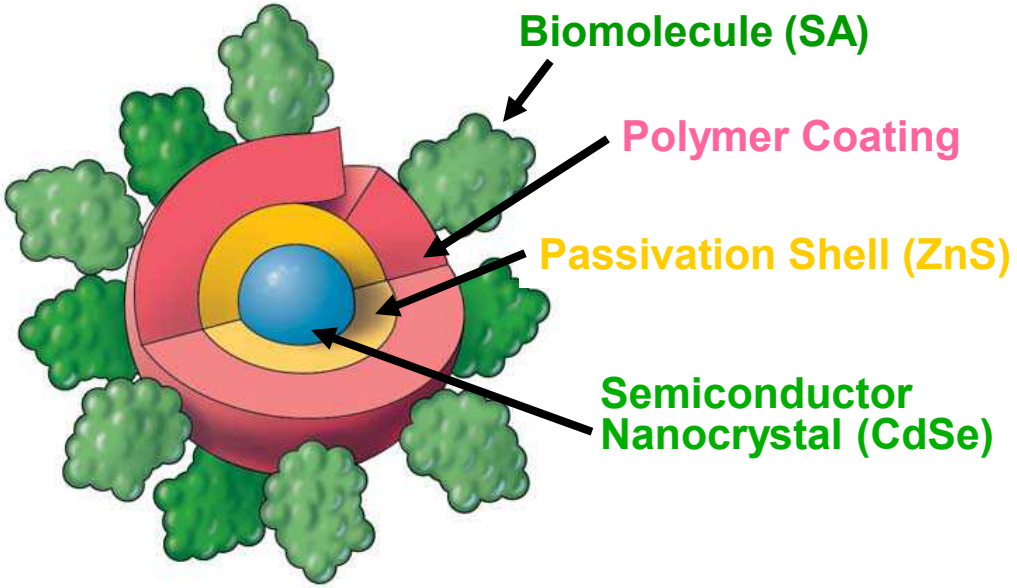
A) No preprocessing
B) Fully preprocessed in automated mode

Sandia National Laboratories

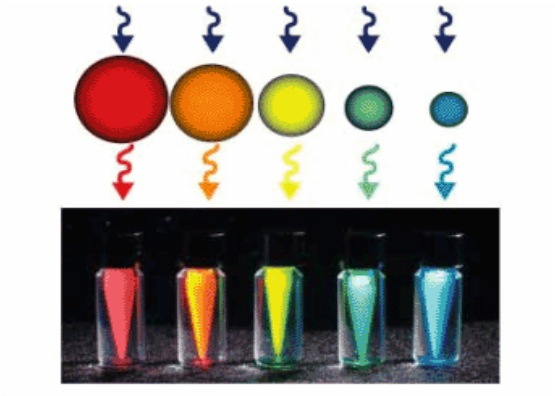
Experimental Example

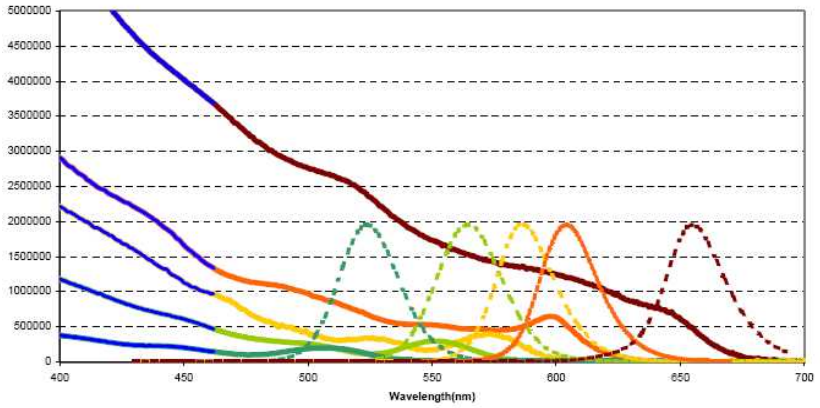
- Quantum dot labeled IgE proteins monitored on membranes of rat basophilic leukemia (RBL) cells
  - 5 QD colors (525, 565, 585, 605, 625, 655 nm)
  - Monitor activation of IgE receptor
  - Time-resolved images (60 frames at 1 frame/4 sec)
  - Data size (100 rows x 104 columns x 60 time points x 512 wavelengths = 640 MB)

Quantum Dots



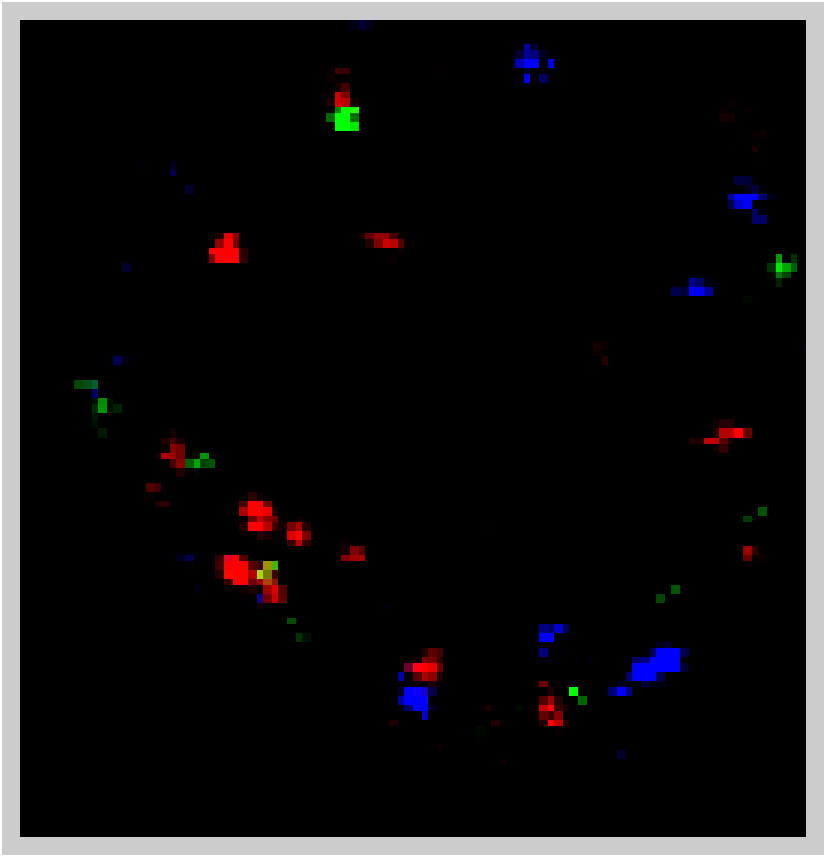
- Broad excitation spectrum
- Narrow emission band
- Brightness
- Photostability
- Flexible bioconjugation
- Electron dense

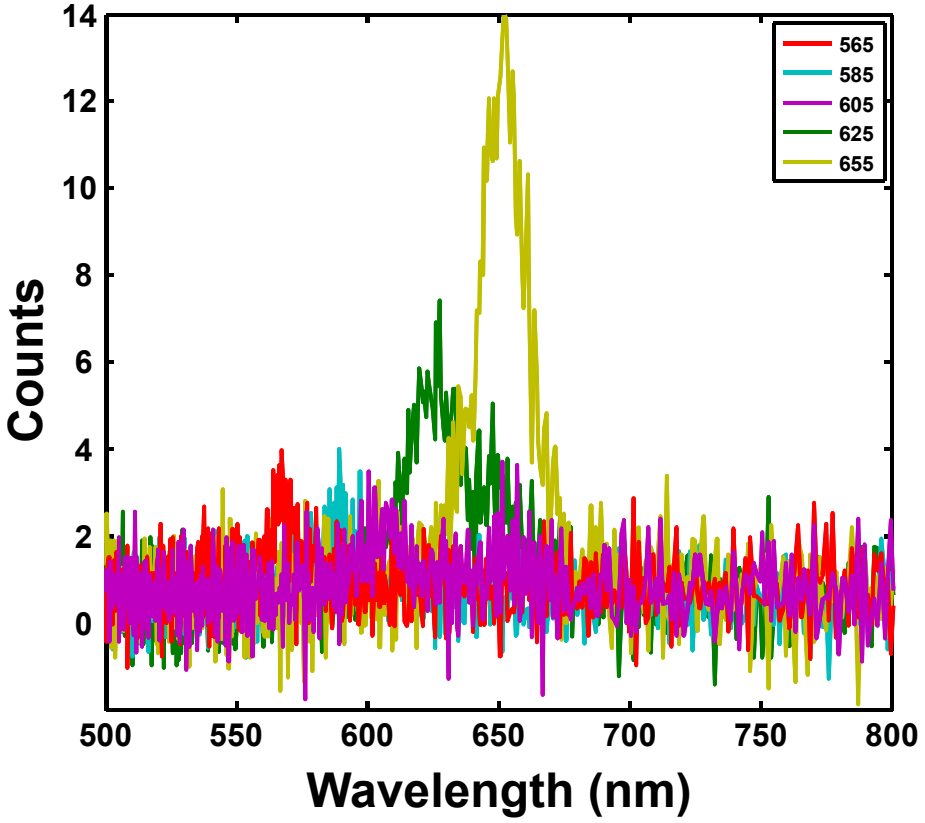




Sandia National Laboratories

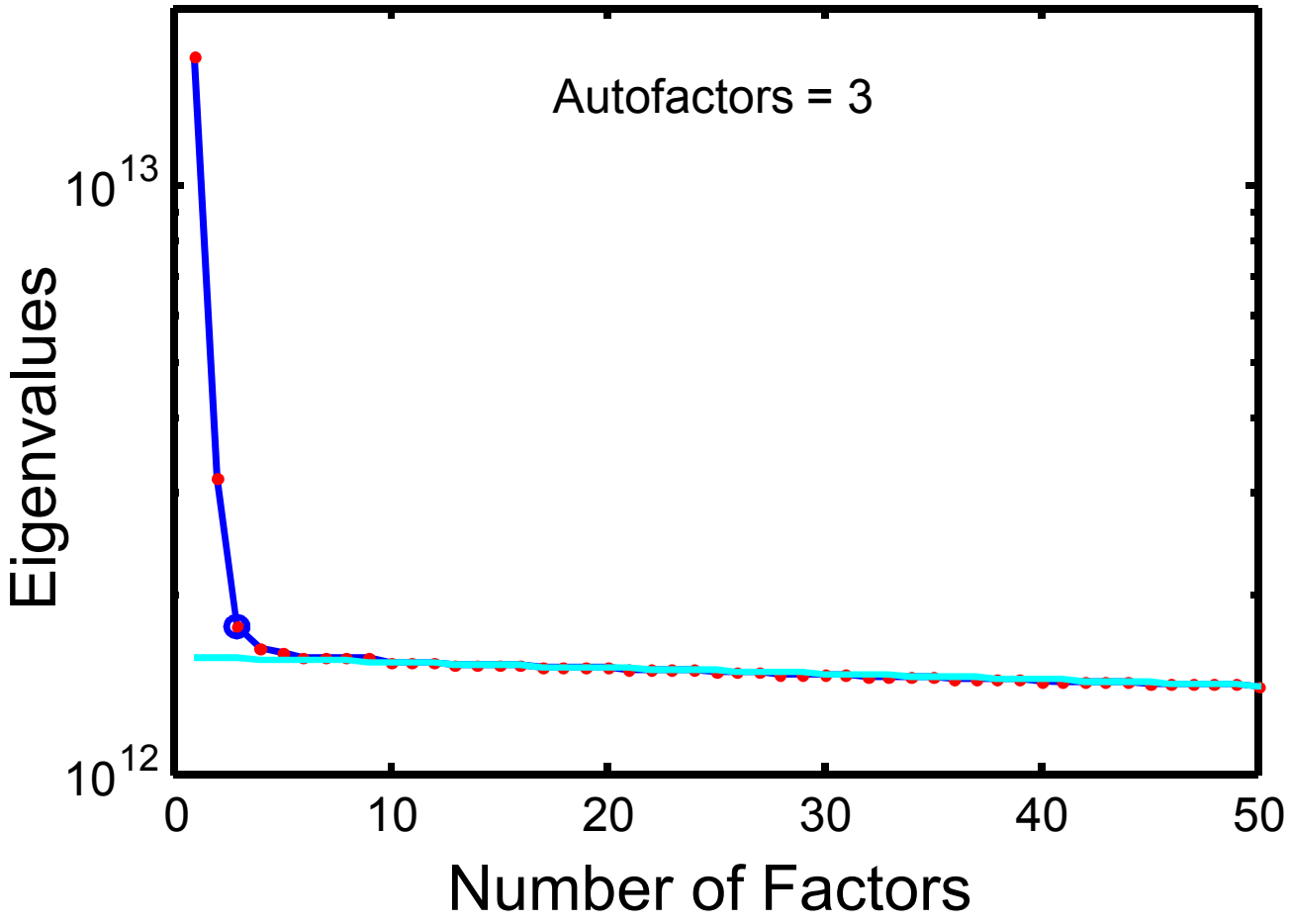
QD Image and Spectra





Sandia National Laboratories

Scree Plot from PCA of Hyperspectral Image of QDs

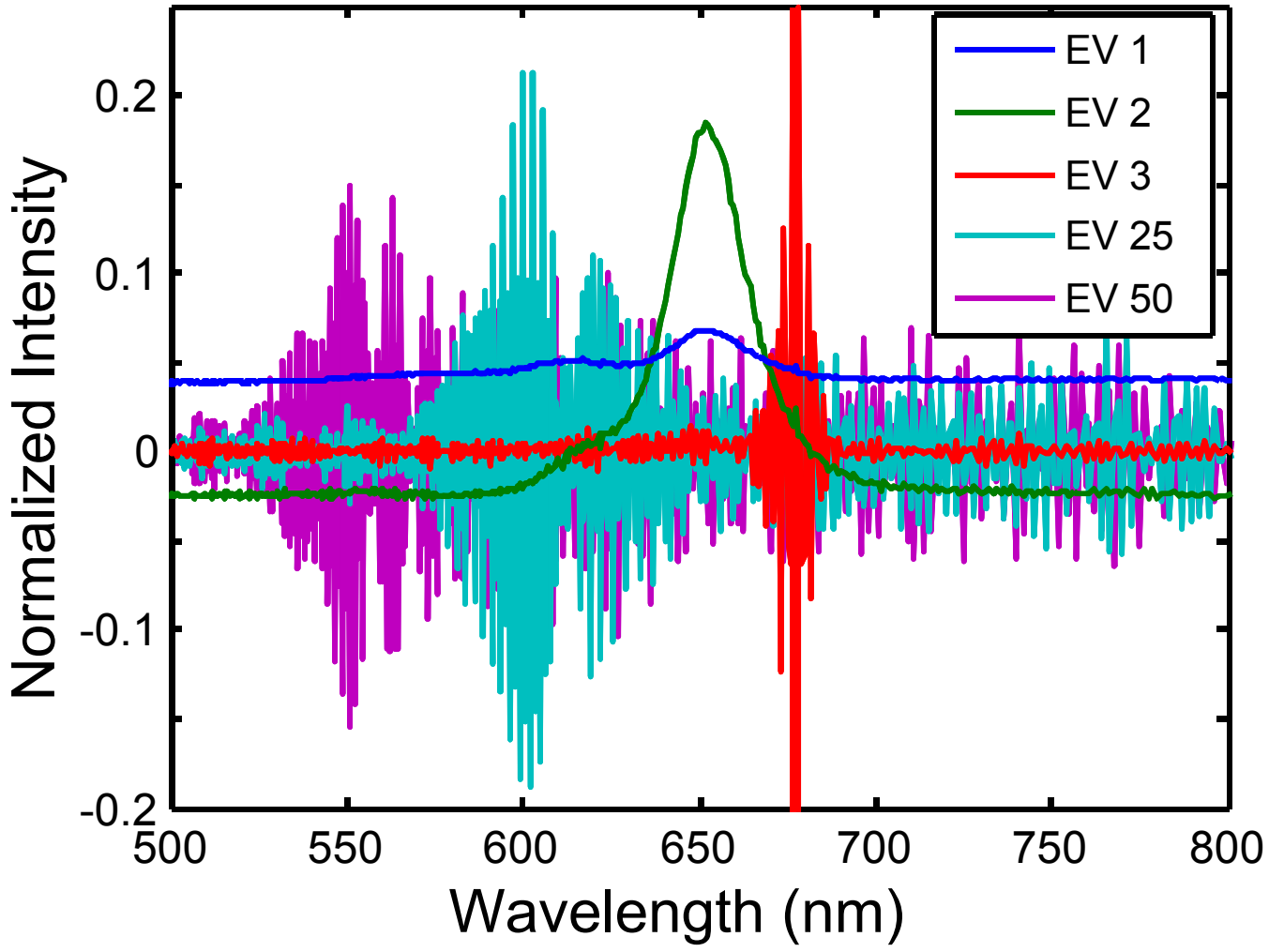


Conventional preprocessing method

1. **Despike**
2. **Subtract dark image**

Sandia National Laboratories

Eigenvectors from QD Hyperspectral Image

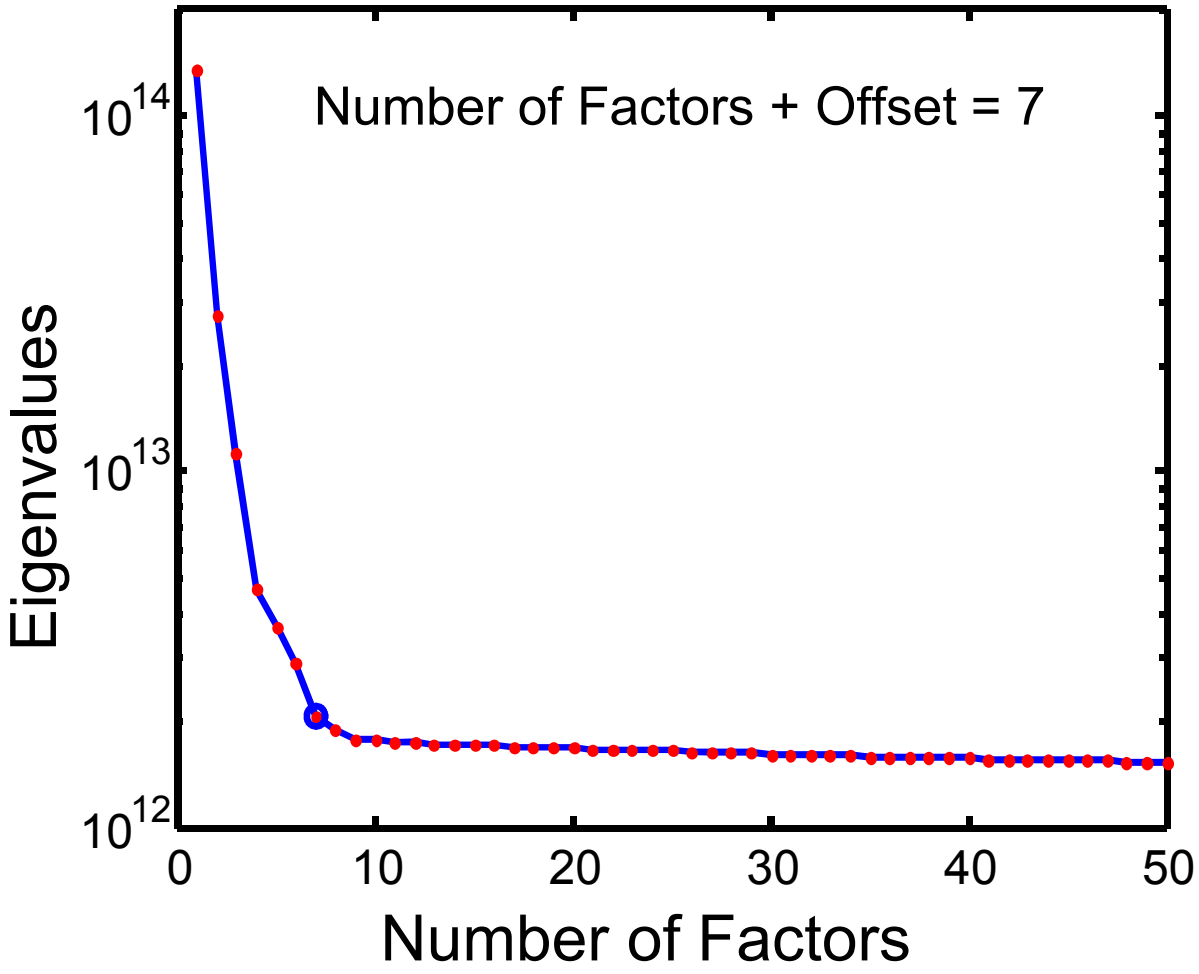


Sandia National Laboratories

Why Does PCA Not Find All the Spectral Components?

- Spectral information from quantum dots is low intensity and present in only a few pixels.
- Correlated read noise errors are in every pixel and represent more variance in the PCA analysis than the variance due to some of the QDs
- Possible solutions:
  - Compress wavelength channels (average successive wavelength pixels) to minimize effects of high frequency correlated spectral noise
  - Compress image spatially (average spatial pixels in x and y) since magnitude of the correlated error is random in the spatial dimensions
  - Apply our full automatic preprocessing of the image
  - Automatic preprocessing is found to be the most sensitive and effective approach

Scree Plot from PCA of 40x Compressed Hyperspectral Image of QDs

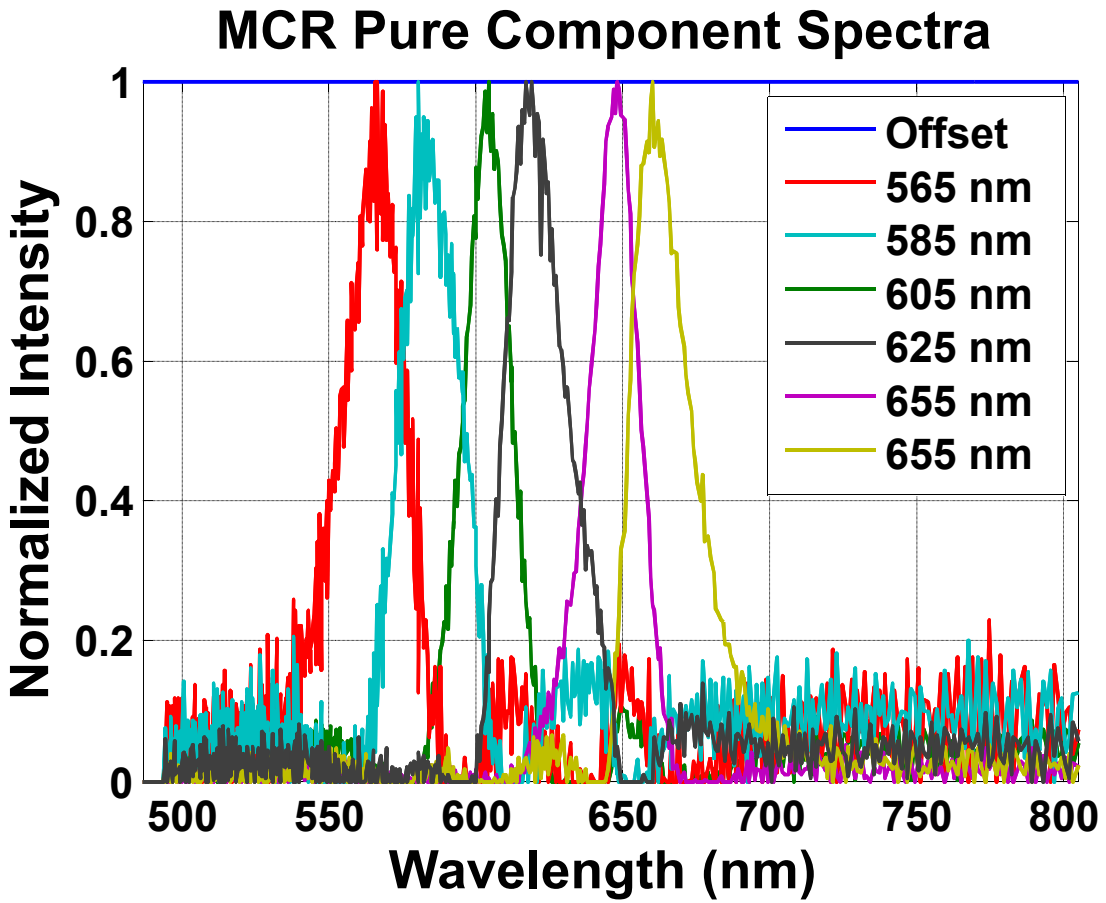


Conventional preprocessing method

1. **Despike**
2. **Subtract dark image**

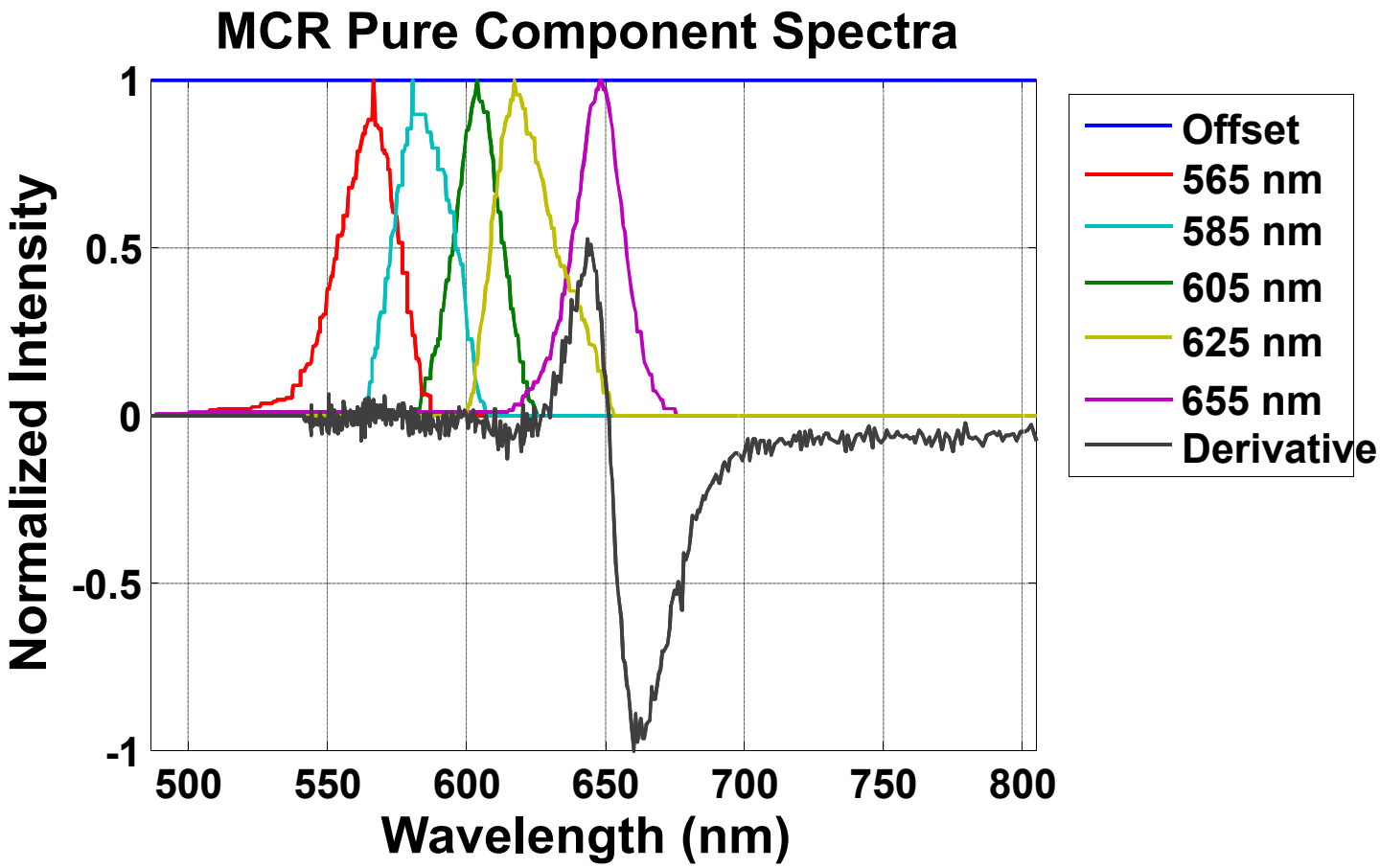
Sandia National Laboratories

MCR Results from 40x Compressed QD Image



Sandia National Laboratories

Alternate MCR Results from 40x Compressed QD Image

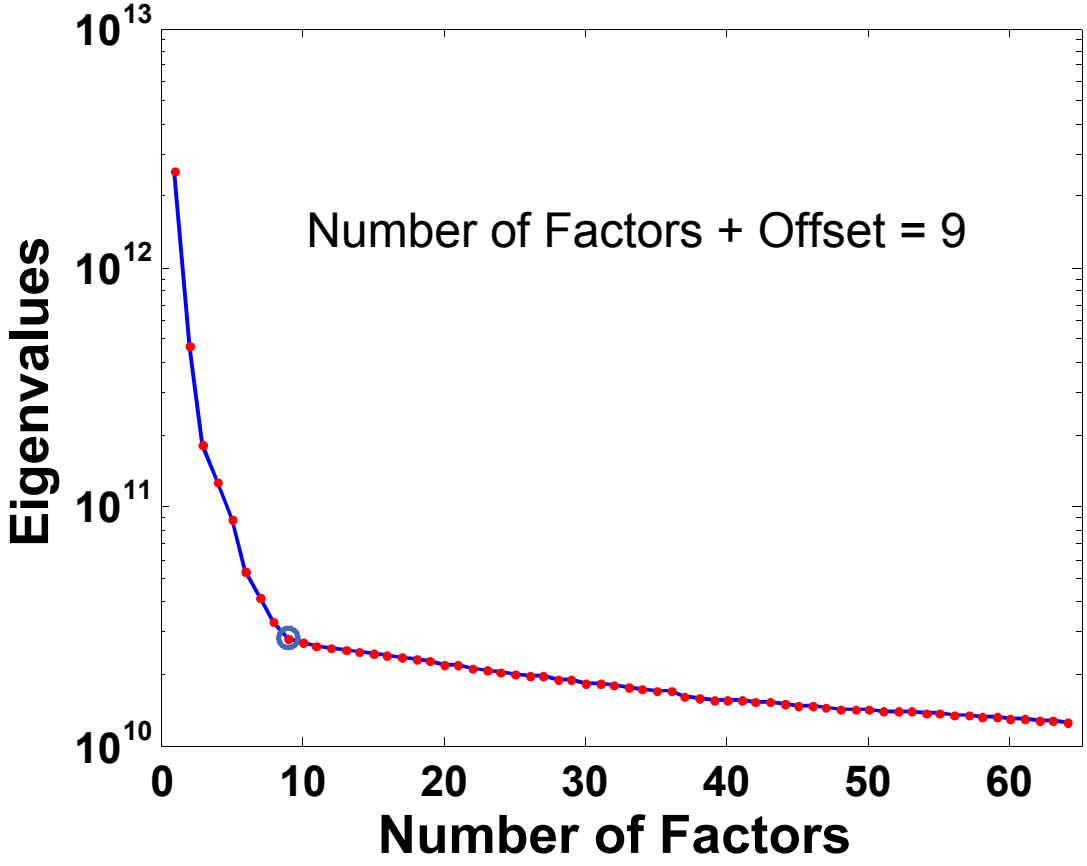


Sandia National Laboratories

Creating Full Spatially-Resolved Image from Compressed MCR Results

- Perform CLS prediction on image spectra using spatially compressed MCR pure-component spectra
- Do not perform PCA on the hyperspectral images (or use all 512 PCA factors)

Scree Plot from PCA of Full Automated Preprocessed QD Image

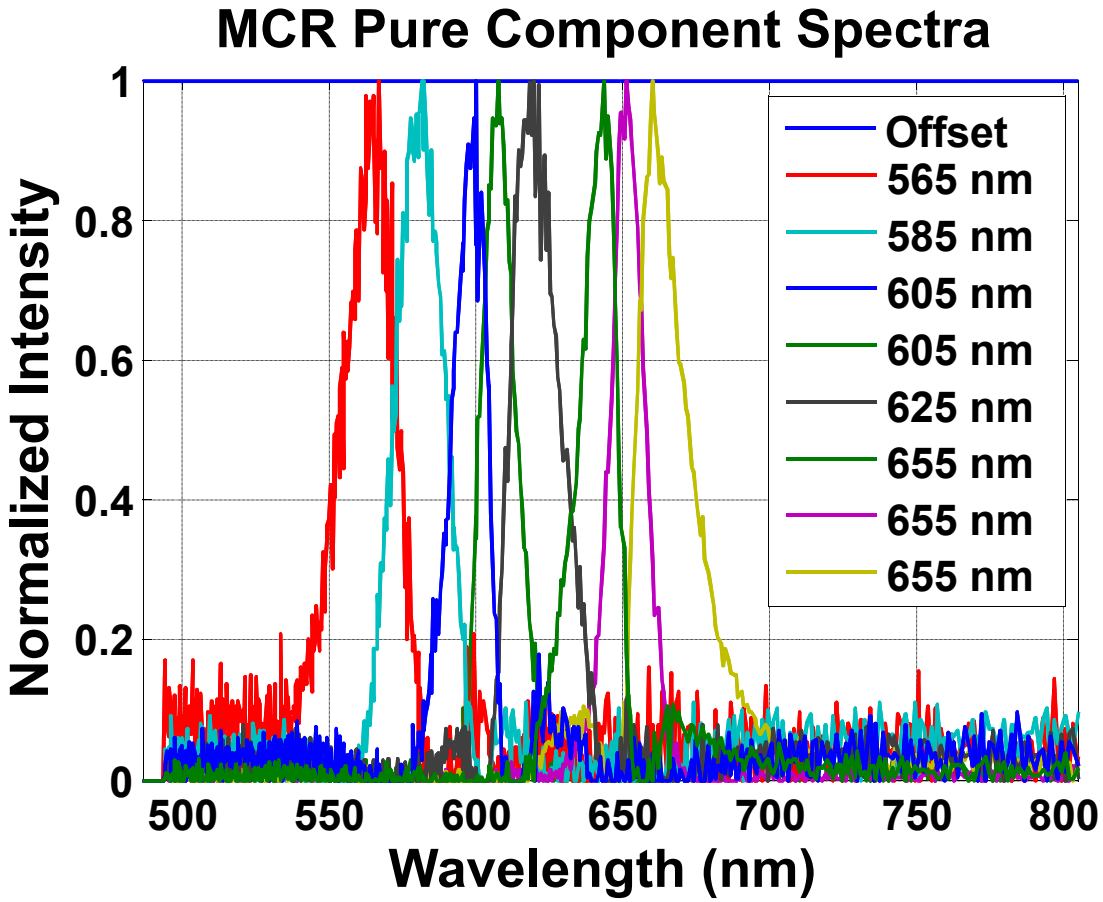


Sandia National Laboratories

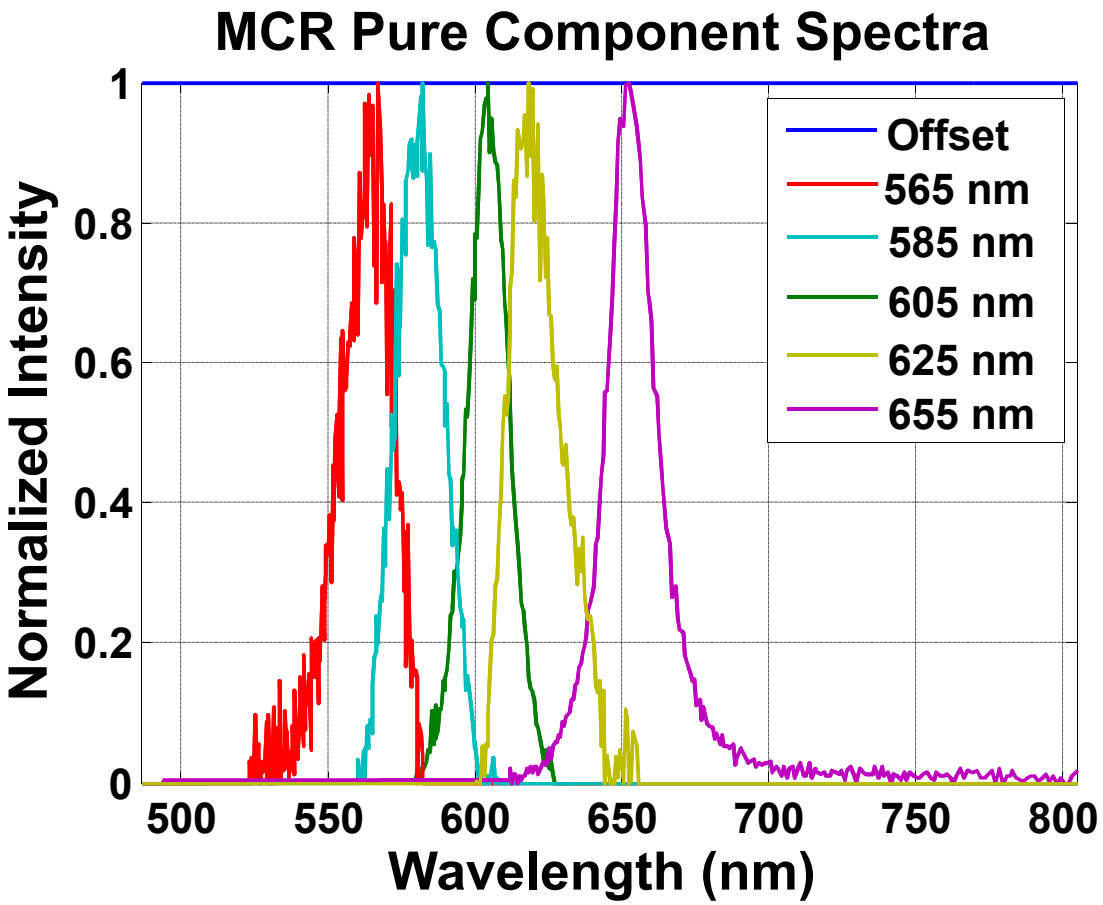
Automatic Preprocessing Applied to the QD Image for MCR Analysis

- Despike hyperspectral image and dark image
- Model and remove structured noise and offset from image using first 20 spectral pixels of the single-scan dark image
- Retain only those pixels whose spectral signal is 3 standard deviations above the noise in the 1st 20 spectral pixels
- Perform MCR on PCA compressed data (64 ev's)
  - Determine number of factors from Scree plot
  - Equality constrain an offset and first zeroed 20 spectral pixels of the random spectral vector starting pure component spectra

MCR Pure Spectra from Fully Preprocessed QD Image

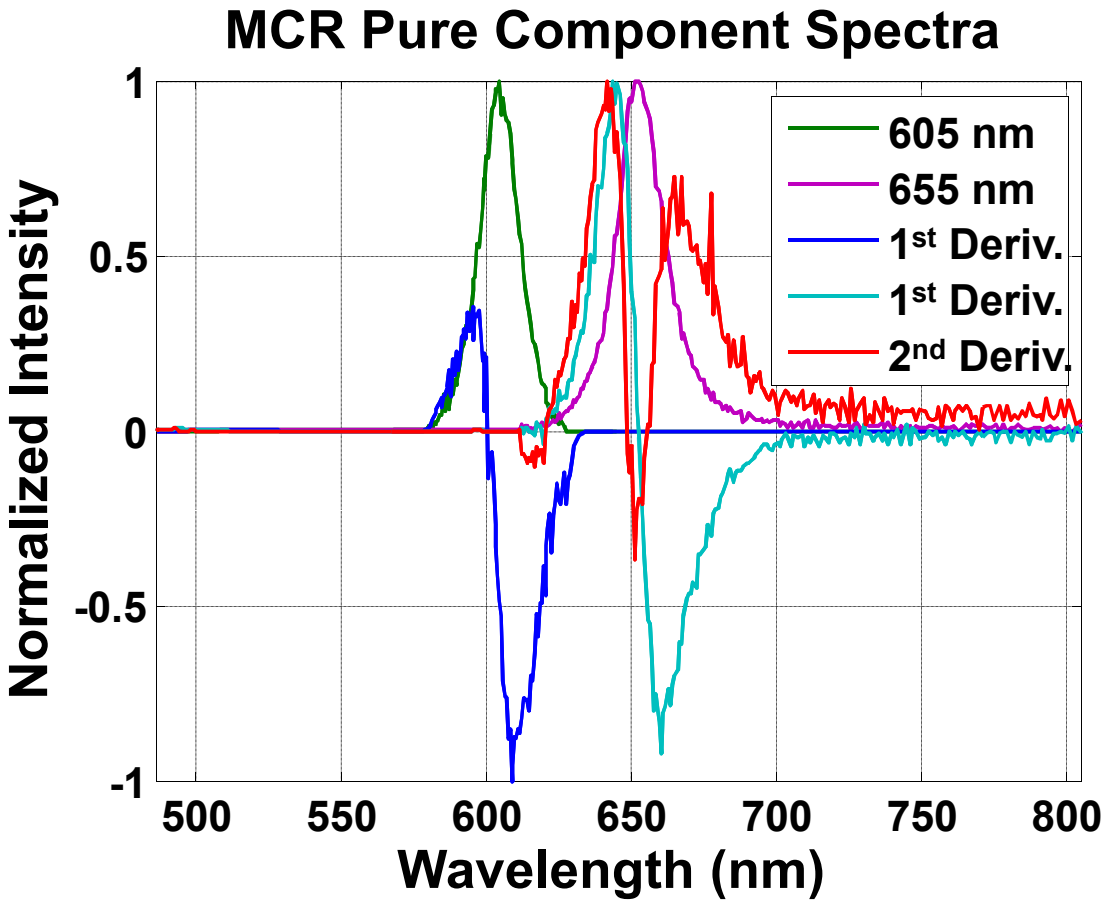


Alternate MCR Results from Fully Preprocessed QD Image (5 QDs)



Sandia National Laboratories

Alternate MCR Results for QD Image (605 nm, 655 nm, derivatives)



Sandia National Laboratories

Conclusions

- Caution: Always examine spectral image data to avoid errors in analyses
- Preprocessing the hyperspectral images is fully automated with Matlab software
- Initial MCR results are fully automated in a batch mode with Matlab software
  - 50 GB of images automatically analyzed in batch mode
- Automation of preprocessing and MCR yields excellent results in the vast majority of analyses
  - <2 min/analysis for 800 MB hyperspectral images
- Experimental approaches to improving MCR analyses
  - Combine multiple images with greater diversity
  - Kinetic photobleaching experiments

Acknowledgements

- Diane Lidke and Nick Andrews, University of New Mexico
- Michael Sinclair, Ryan Davis, Mark Van Benthem, and Mike Keenan, Sandia National Laboratories

Sandia National Laboratories

Sandia National Laboratories