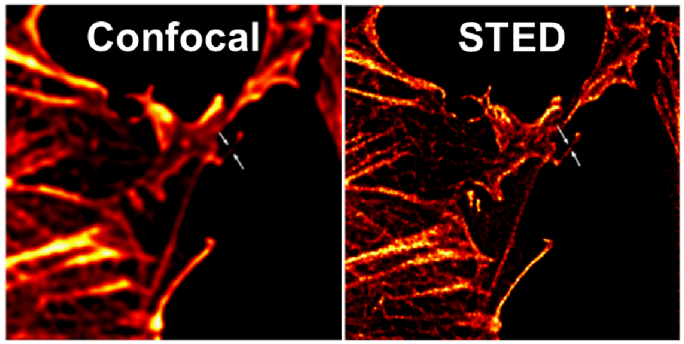
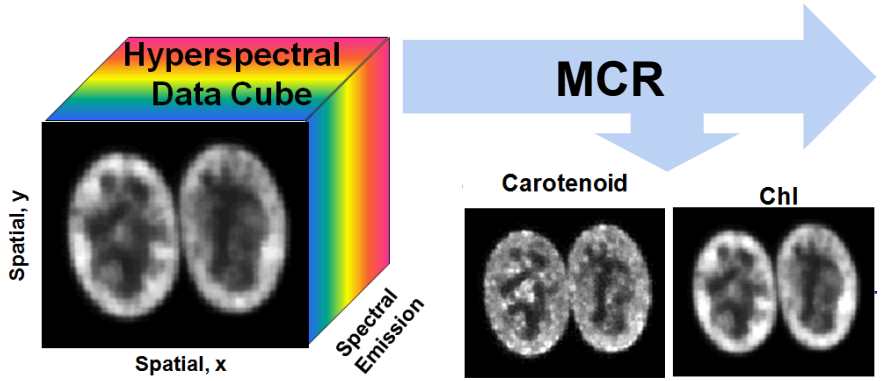
SAND2017-3611C



Adapted from Howard Vindin - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=40722030



Hyperspectral Super-Resolution Imaging and Data Analysis

Stephen M. Anthony

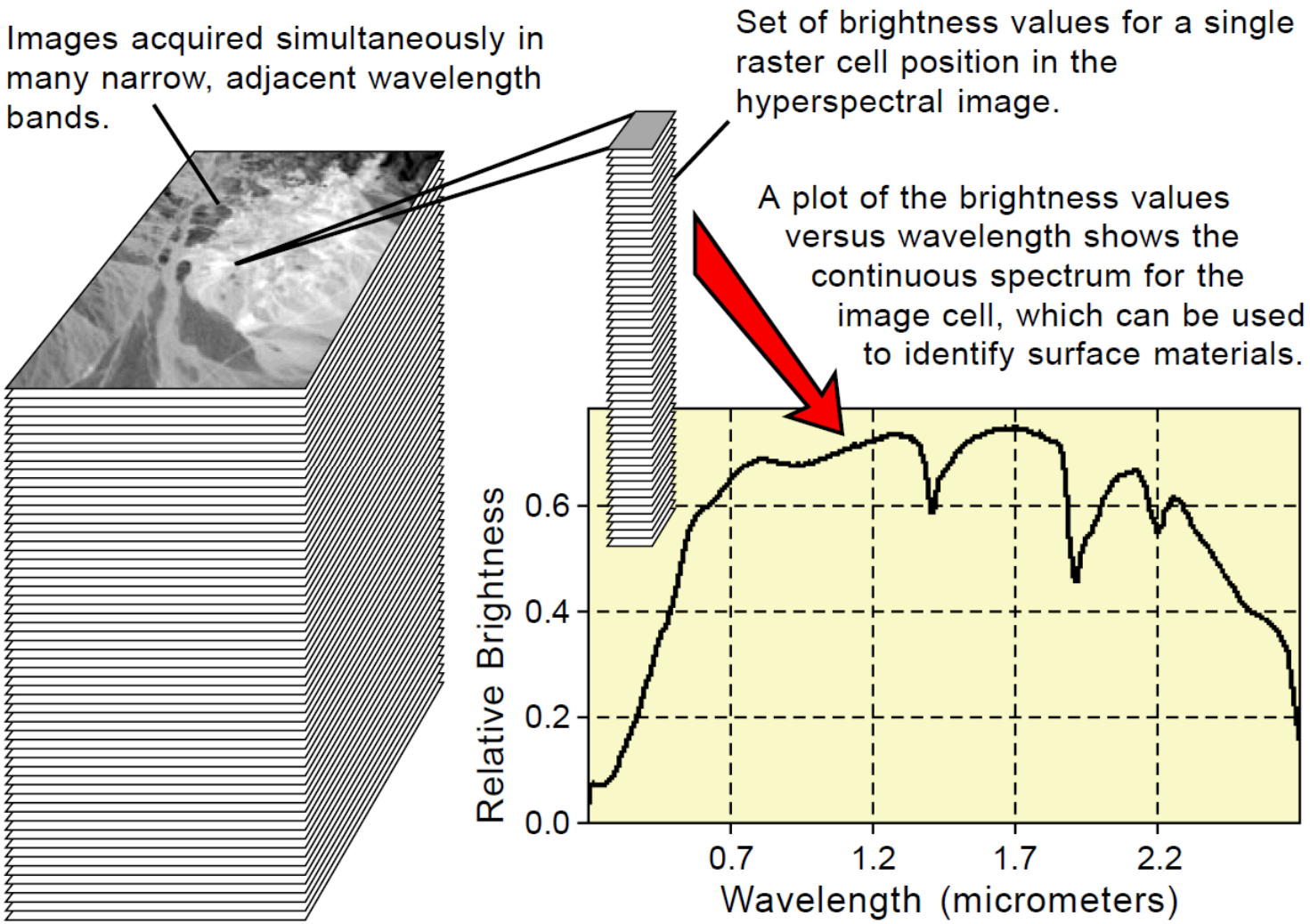
U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

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. SAND NO. 2011-XXXXP

Outline

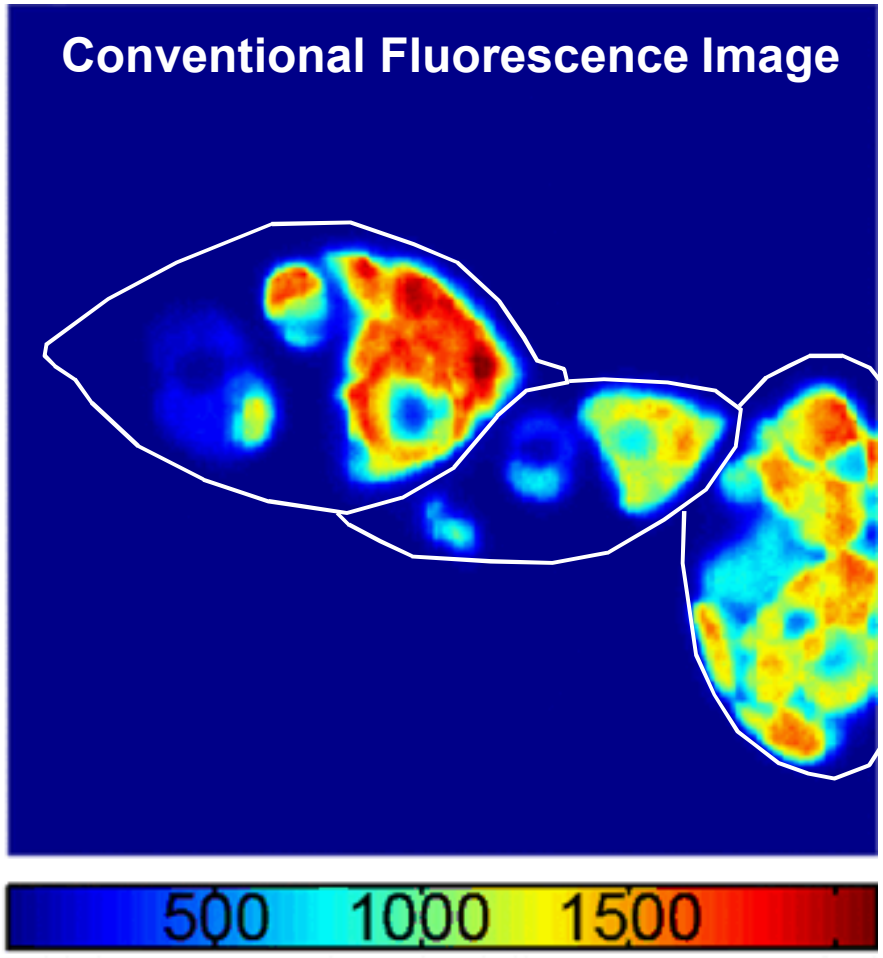
- **Why hyperspectral?** – What is hyperspectral imaging and what are its benefits?
- **Sandia's hyperspectral microscopes** – What are the systems I typically work with?
- **Analyzing hyperspectral data** – What can multivariate curve resolution (MCR) do for you?
- **Improving MCR** – Ongoing work to improve its capabilities.
- **Trilinear Data** – How to leverage additional information.

Introduction to Hyperspectral Imaging



Smith, R. B. (2012) Introduction to Hyperspectral Imaging. Microimages
http://www.microimages.com/documentation/Tutorials/hyprspec.pdf

Why Use Hyperspectral Imaging?

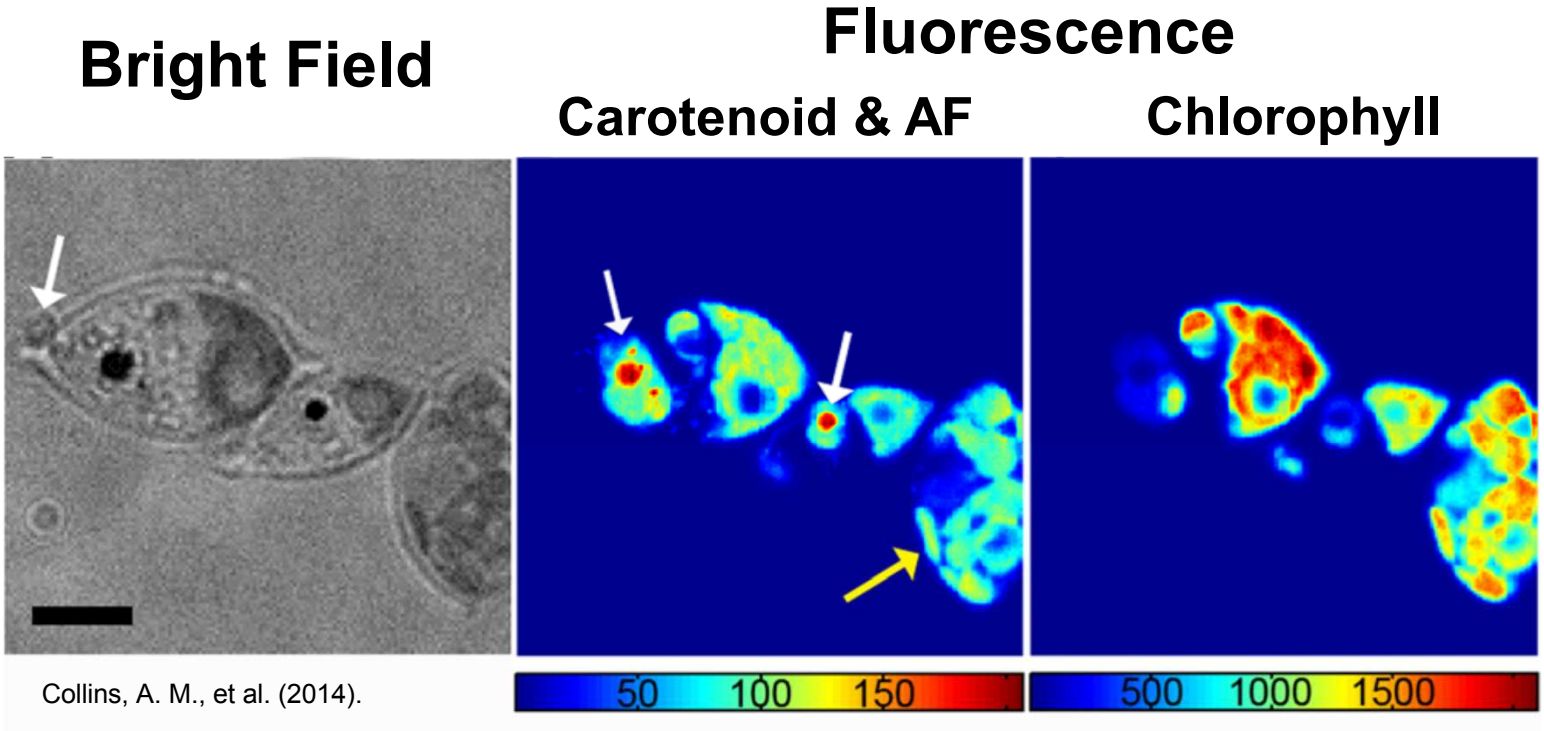


Adapted from Collins, A. M., et al. (2014). "Host Cell Pigmentation in Scenedesmus dimorphus as a Beacon for Nascent Parasite Infection." Biotechnology and Bioengineering **111(9): 1748-1757.**

Image of the endogenous fluorescence from *S. dimorphous* (algae) undergoing parasitic infection by *A. protococcarum*.

- Approximate cell borders are hand-drawn in white.
- Two of the cells contain parasitic vacuoles.
- **Can you spot the parasitic vacuoles?**

Why Use Hyperspectral Imaging?

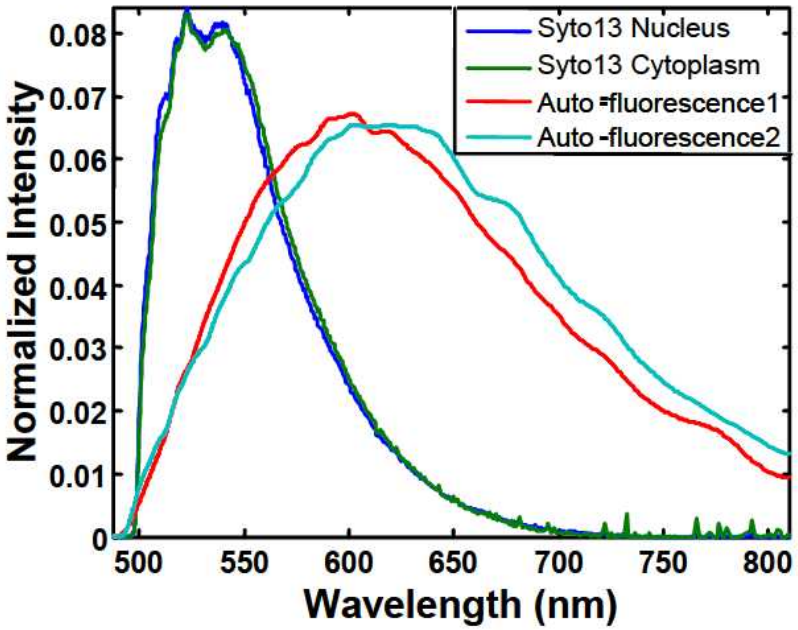


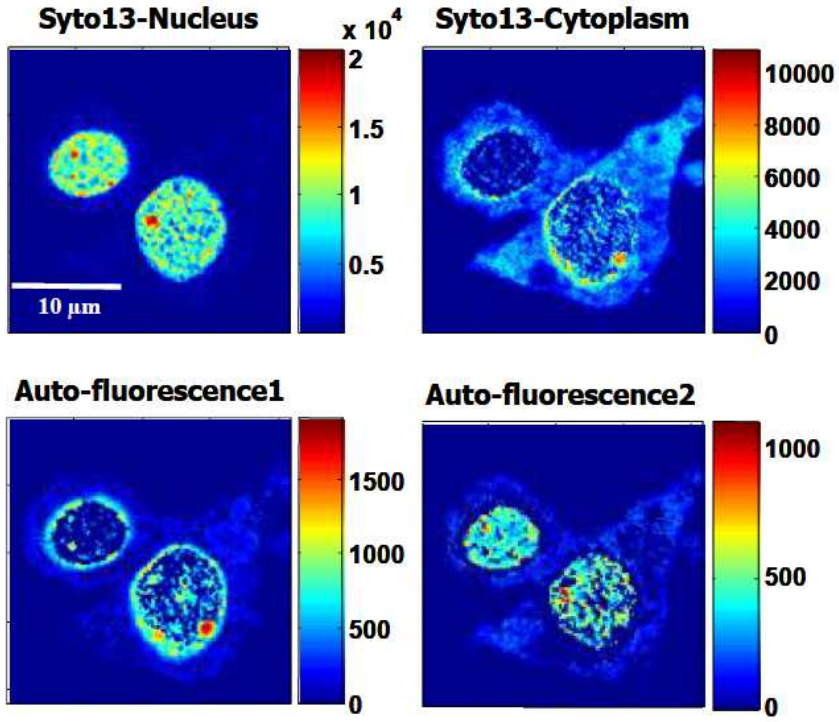
Collins, A. M., et al. (2014).

- Parasitic vacuoles (white arrows) are easily spotted using the combined carotenoid and autofluorescence signal.
- Spotting them is nearly impossible when examining all the fluorescence together as the chlorophyll signal dominates.
- **Hyperspectral imaging reveals otherwise hidden features.**

Why multispectral is not enough

Haaland, D. M., et al. (2007). "Hyperspectral confocal fluorescence imaging of cells" Next-Generation Spectroscopic Technologies **6765: 76509-76509.**





Left) Fluorescence spectra for two Syto 13 and two autofluorescence emission components. Right) Relative concentration of the components' spatial distributions in mouse macrophage cells (Raw 264.7).

- Multispectral imaging (e.g. filter-based microscopes) would only distinguish Syto 13 from autofluorescence – two components.
- **Hyperspectral imaging can distinguish nearly identical spectra.**

Hyperspectral Imaging Applications

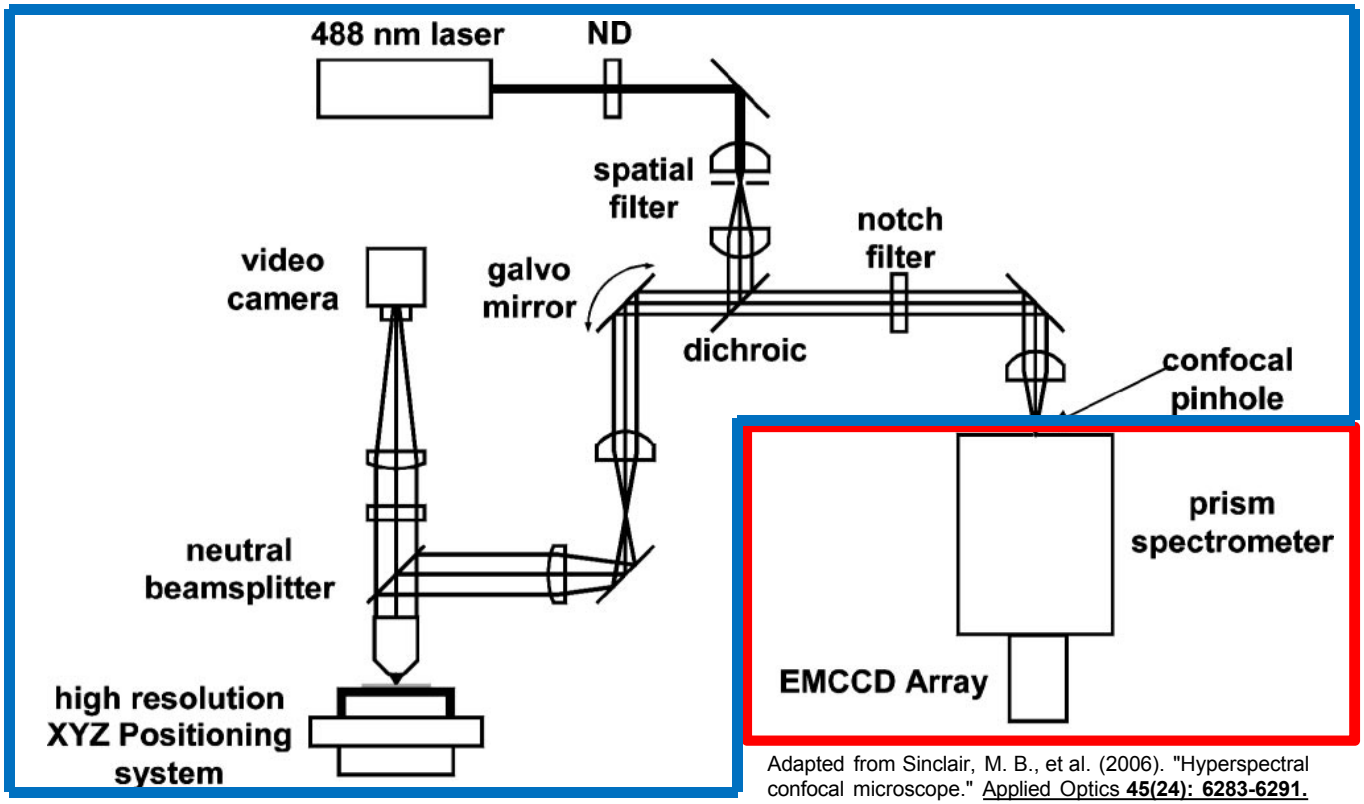
Sandia National Laboratories

- Fluorescence or Raman microscopy – cell signaling
- Agriculture (satellite or drone-based) – monitoring crop locations (poppy fields), crop health, disease outbreak
- Chemical detection – Detect airborne chemical hazards at ppm levels up to 5 km away
- Mineralogy – disturbed ground indicative of improvised explosive devices

Outline

- **Why hyperspectral?** – What is hyperspectral imaging and what are its benefits?
- **Sandia's hyperspectral microscopes** – What are the systems I typically work with?
- **Analyzing hyperspectral data** – What can multivariate curve resolution (MCR) do for you?
- **Improving MCR** – Ongoing work to improve its capabilities.
- **Trilinear Data** – How to leverage additional information.

How to Build a Hyperspectral Microscope



Adapted from Sinclair, M. B., et al. (2006). "Hyperspectral confocal microscope." Applied Optics **45(24): 6283-6291.**

Schematic diagram of Sandia's hyperspectral confocal microscope

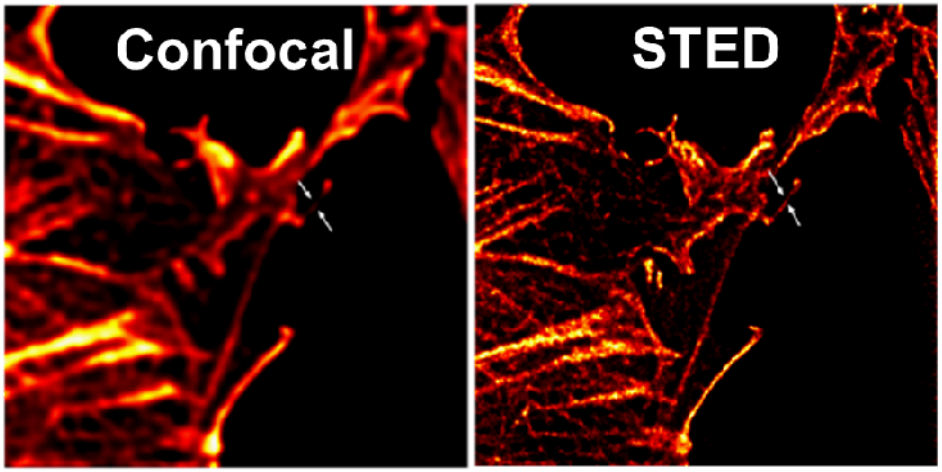
Hyperspectral Confocal Microscope =

Confocal Microscope + Spectrometer

Hyperspectral STED Microscope

Sandia National Laboratories

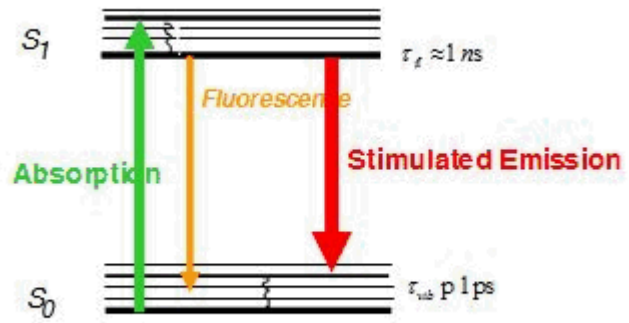
- Stimulated Emission Depletion (STED) is a super-resolution microscopy technique
  - Super-resolution microscopy won the 2014 Nobel prize in chemistry
- STED dramatically improves the spatial resolution (~30 nm typical, <3 nm reported)



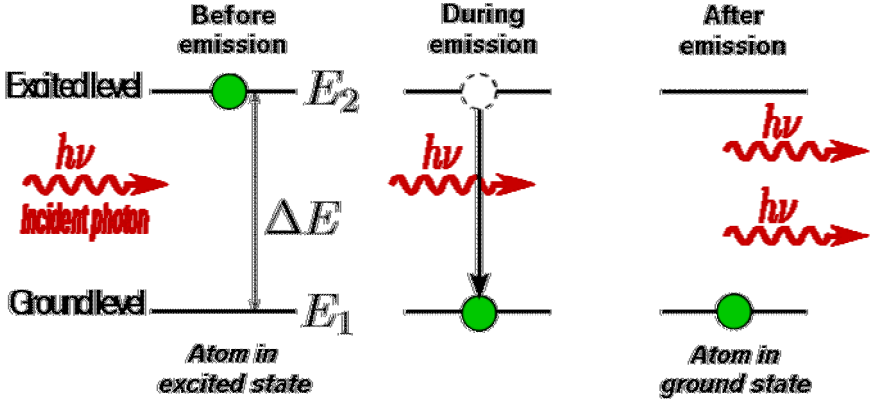
- Building world's first hyperspectral STED microscope
  - Patented by Jeri Timlin (8631) and Jesse Aaron

Stimulated Emission Details

Basic Principle of Stimulated Emission

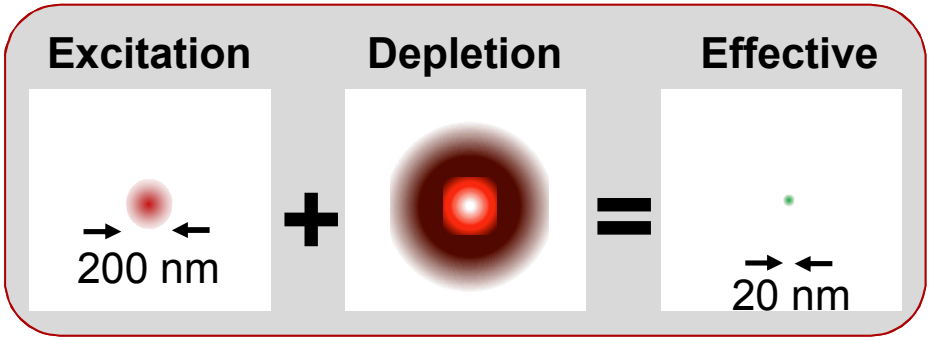


https://en.wikipedia.org/wiki/File:STED_Jablonski.jpg



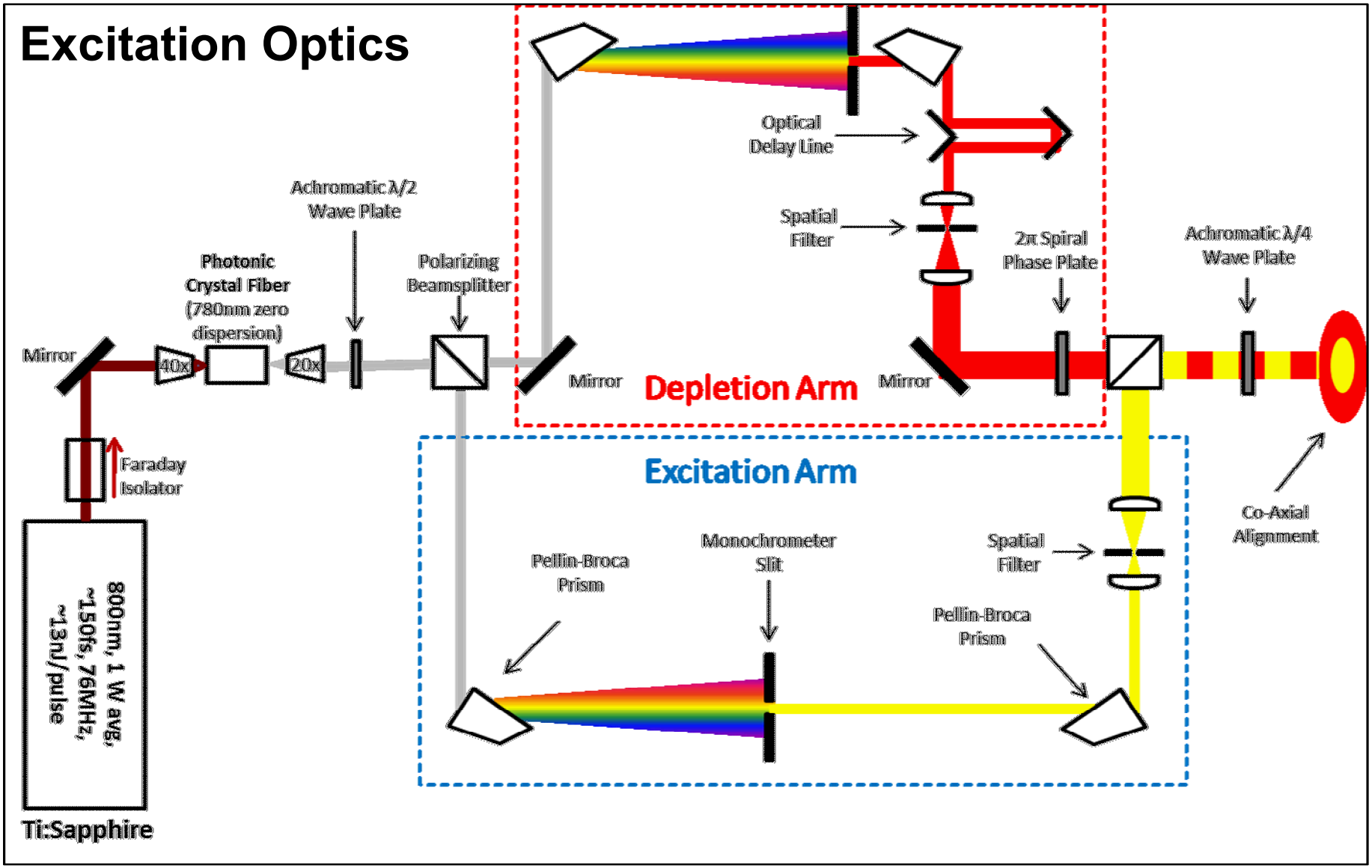
By V1adis1av - Own work, GFDL, https://commons.wikimedia.org/w/index.php?curid=3983414

Generating the STED Point Spread Function (PSF)



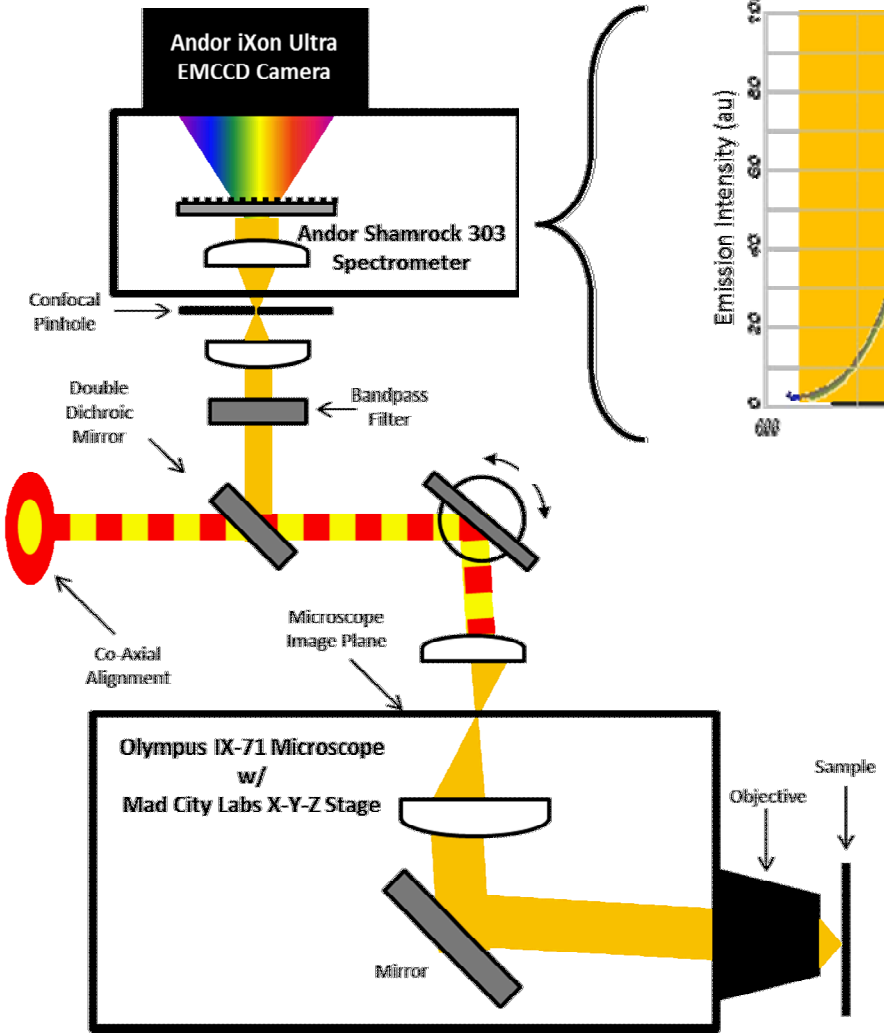
Neither beam PSF can exceed the diffraction limit, but the effective PSF can!

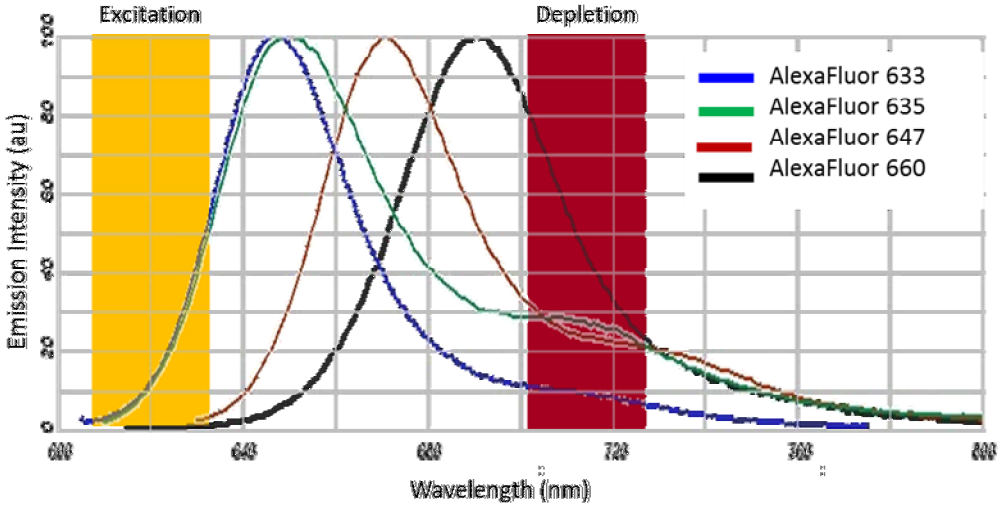
Building a Hyperspectral STED



Building a Hyperspectral STED

Detection Optics





Design Considerations

- Tunable wavelength for both excitation and depletion beams
- Can be optimized for any STED fluorophore with exchange of a single optic (the dichroic)

Outline

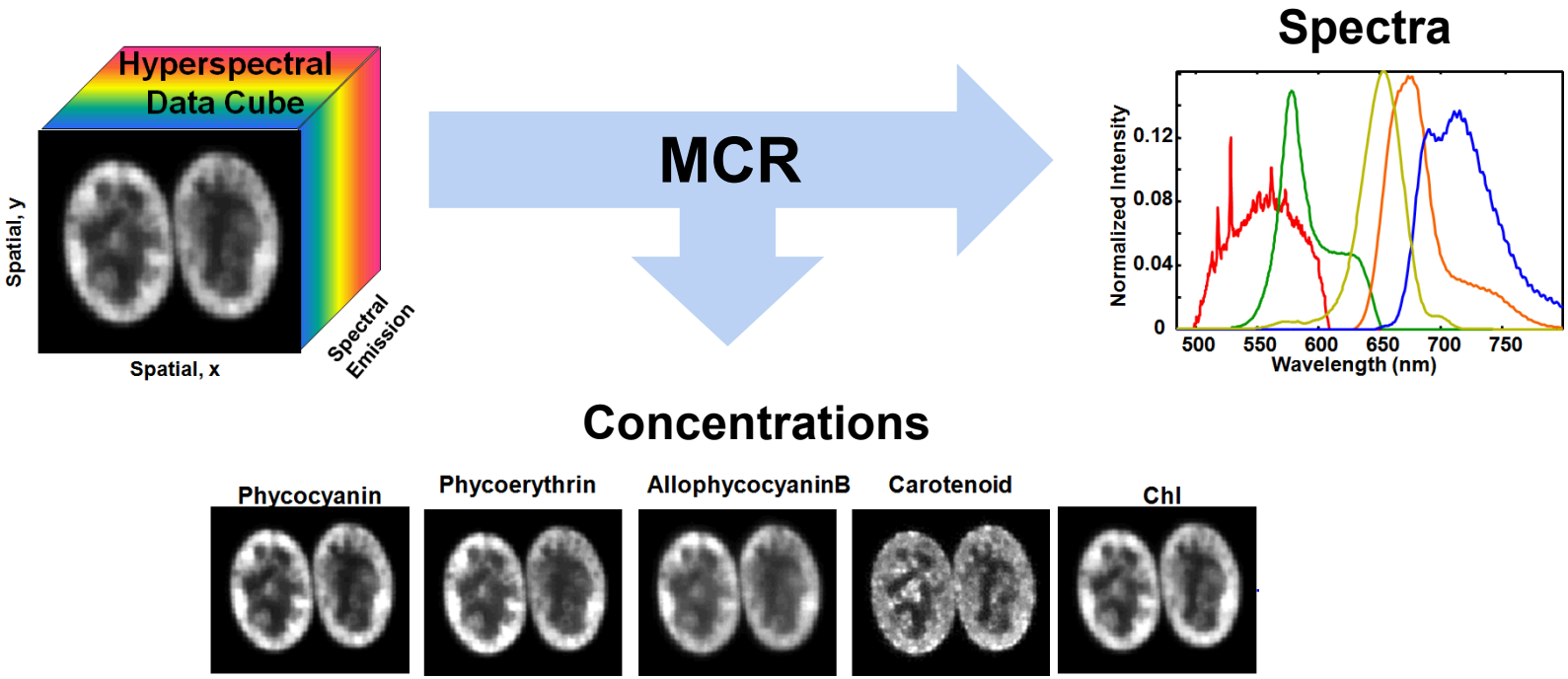
- **Why hyperspectral?** – What is hyperspectral imaging and what are its benefits?
- **Sandia's hyperspectral microscopes** – What are the systems I typically work with?
- **Analyzing hyperspectral data** – What can multivariate curve resolution (MCR) do for you?
- **Improving MCR** – Ongoing work to improve its capabilities.
- **Trilinear Data** – How to leverage additional information.

Multivariate Curve Resolution (MCR)

Chemometric factor analysis (such as MCR) extracts:

1) the spectra of the pure compounds &

2) their relative concentrations at each position.



Common Spectral Image Analysis Methods

Sandia National Laboratories

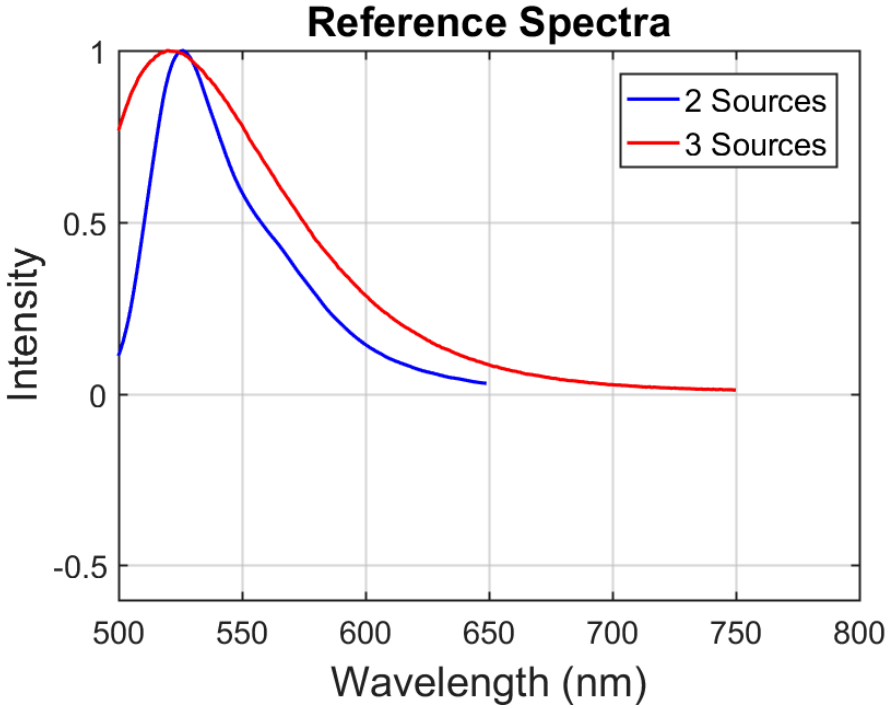
Required knowledge of system behavior

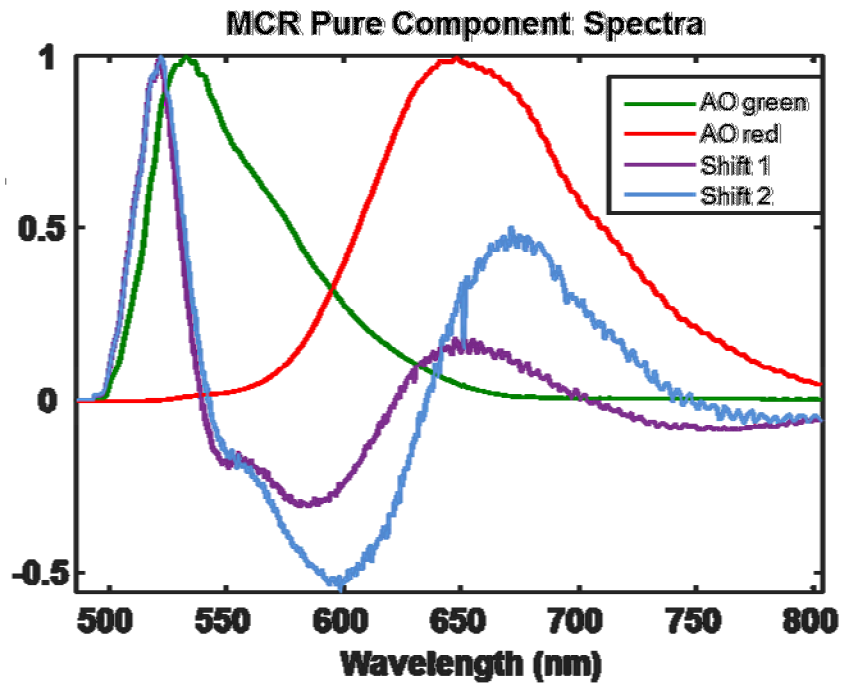
Mathematical Complexity, overall accuracy

- Univariate methods
  - → Band integration, peak height, peak positions
  - → **Isolated bands, no spectral interference**
- Multivariate methods
  - Unmixing methods
    - → Least squares prediction based (e.g. classical least squares)
    - → ***A priori* knowledge**
      - → Spectral shapes or pure image pixels
  - Factor Analysis methods
    - → Principle components analysis (PCA) , Factor analysis, SIMPLISMA, self modeling curve resolution, multivariate curve resolution **(MCR)**
    - → Data defines
    - → **No *a priori* knowledge** of spectral shapes/pure pixels
      - → Need number of components
      - → Constraints to narrow solution space

Why Aren't the Spectra Known?

"Acridine Orange is a cell-permeant nucleic acid binding dye that emits **green fluorescence** when bound to dsDNA and **red fluorescence** when bound to ssDNA or RNA." - ThermoFisher Scientific





Reference spectra are not always available, and when available do not always capture the complete spectral properties.

Broadly Applicable

Sandia National Laboratories

Works for any data satisfying the linear additive model:

$$D = CS^T + E$$

- **Fluorescence spectroscopy**
- **Raman spectroscopy**
- **Mass spectroscopy**
- Infrared satellite imagery

MCR Assumptions

1. **Assumes a linear additive model:**

$$\mathbf{D = CS^T + E}$$

D = Data matrix — nPoints X nWavelengths

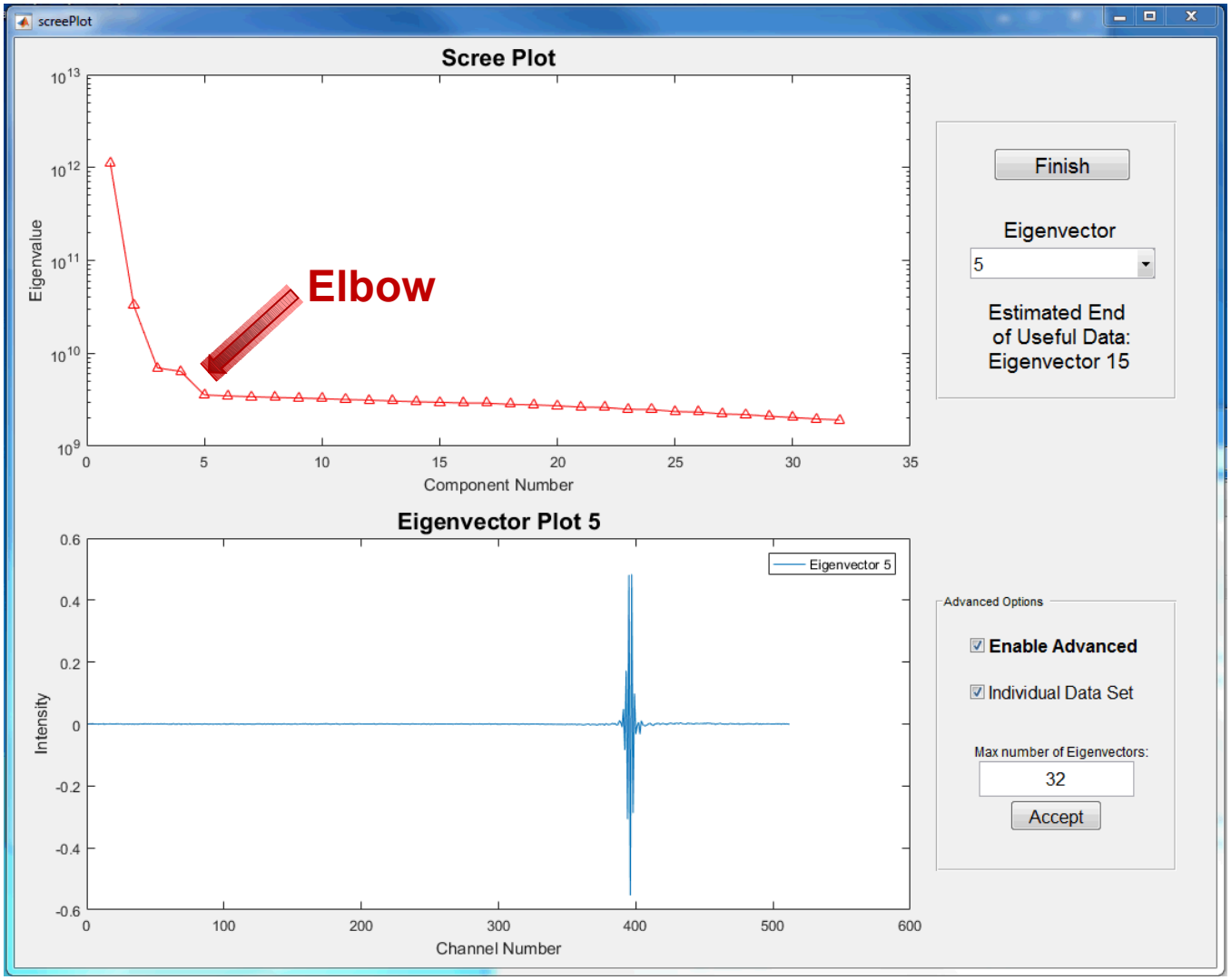
C = Concentrations matrix — nPoints X nComponents

S^T = (Spectra matrix)^{Transpose} — nComponents X nWavelenths

E = Noise (error) matrix — nPoints X nWavelengths

2. The # of spectral components is known or can be estimated

Determining the # of Spectra



Scree plots allow rough estimation of the number of spectra.

Elbow (transition to flat) in scree and eigenvectors that look like noise indicate no more components.

Basic Operation

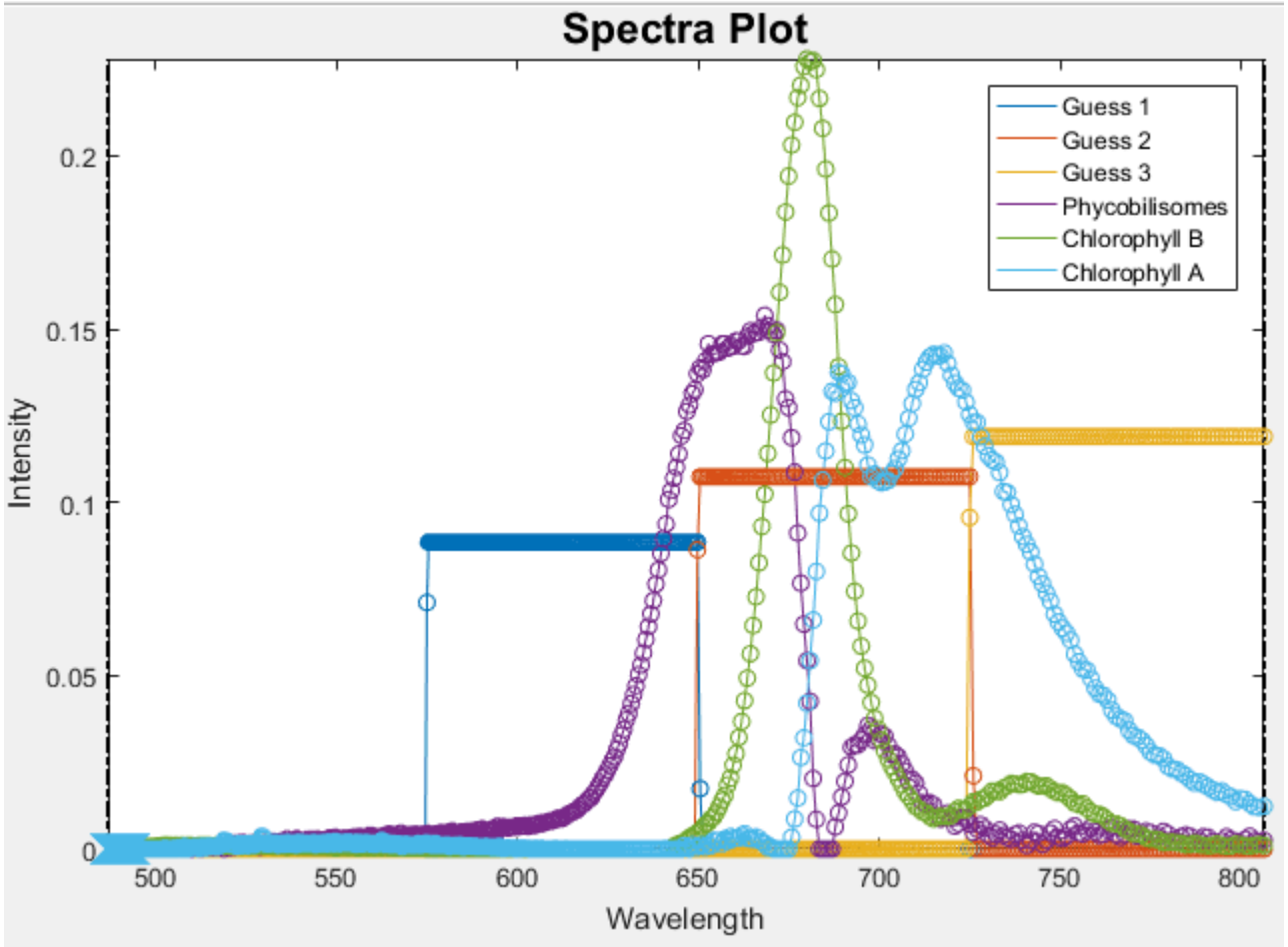
$$D = CS^T + E$$

- D is known
  - If C were known, could solve for S
  - If S were known, could solve for C
- Constrained Alternating Least Squares
  1. Provide an initial guess for S (or C)
  2. Solve for C based upon current S guess, enforcing constraints
  3. Solve for S based upon current C guess, enforcing constraints
  4. Repeat steps 2 & 3
  5. **Iterations converge on solution**

Advantages of MCR

- Extracts underlying relationships from complex data sets
- No *a priori* knowledge needed
- Signals below the noise level can be detected!
- Physically meaningful constraints
  - Negative concentrations not allowed
  - Negative intensities not allowed
- Efficient algorithms developed at Sandia
  - Keenan, M. R. and P. G. Kotula (2003). Apparatus and system for multivariate spectral analysis, Google Patents.

No *A Priori* Spectra Required



Hyperspectral Confocal Fluorescence Microscopy of R61 *Acaryochloris*

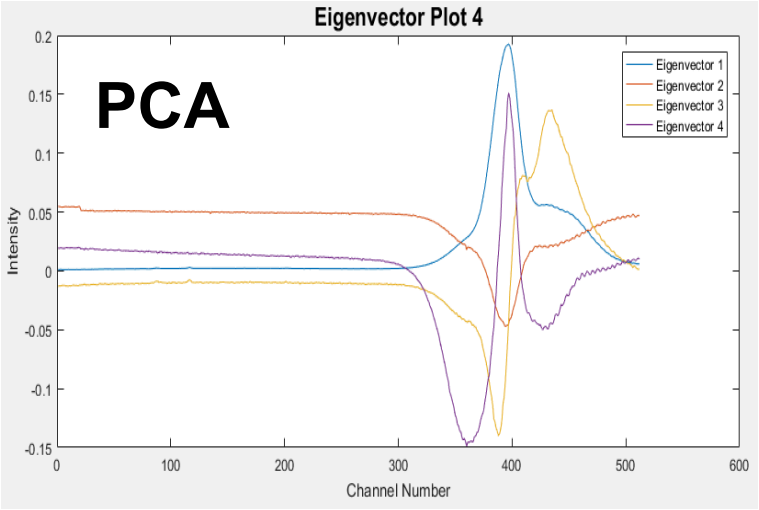
Even when initialized with naïve guesses,

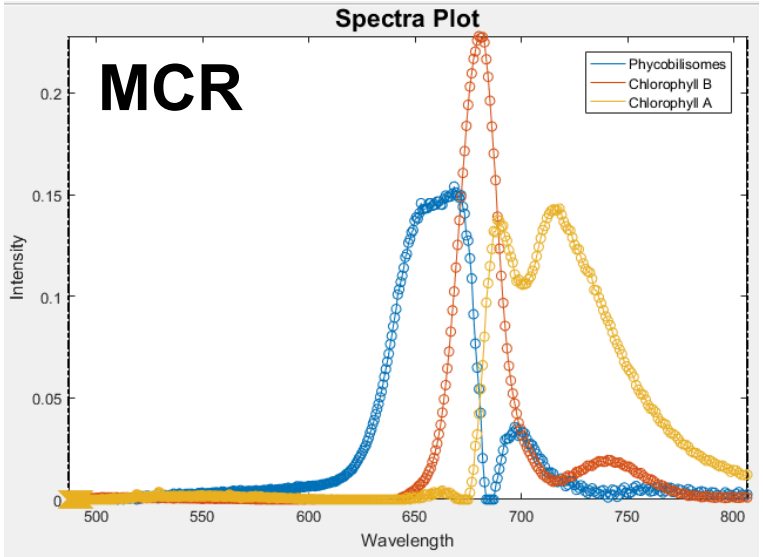
MCR determines the spectrum of chlorophyll B and localizes where it is present.

Why MCR vs. PCA?

- Three related factor analysis methods
  - Multivariate Curve Resolution (MCR)
  - Principal Component Analysis (PCA)
  - Independent Component Analysis (ICA)
- All resolve the data into pure spectral components and concentrations without a priori information
- **Different Constraints**
  - MCR – Physical and Chemical Constraints (e.g. no negative concentrations, no negative intensities)
  - PCA – Linearly uncorrelated
  - ICA – Statistically Independent

Why MCR vs. PCA?



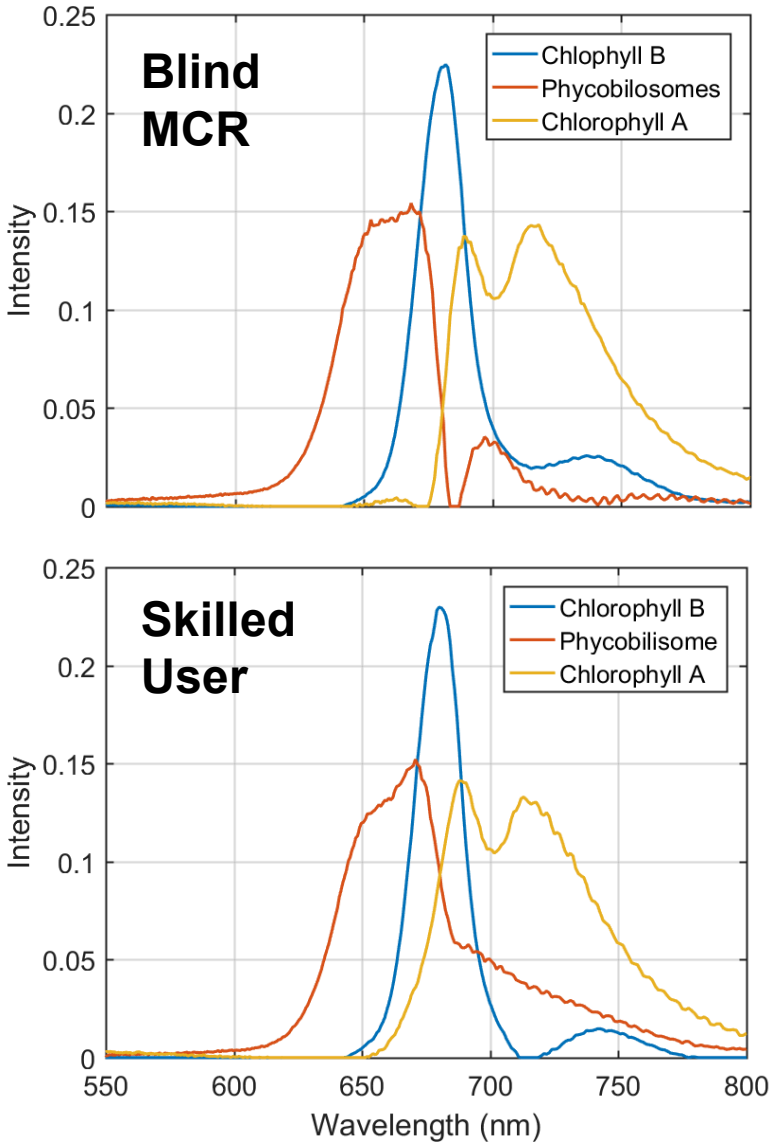


- PCA
  - Recovers 4 spectral components when only 3 are present
  - Eigenvectors do not look like normal spectra – major negative portions
- MCR
  - Recovers the correct number of components
  - Components generally correspond to actual spectra

Outline

- **Why hyperspectral?** – What is hyperspectral imaging and what are its benefits?
- **Sandia's hyperspectral microscopes** – What are the systems I typically work with?
- **Analyzing hyperspectral data** – What can multivariate curve resolution (MCR) do for you?
- **Improving MCR** – Ongoing work to improve its capabilities.
- **Trilinear Data** – How to leverage additional information.

The Art of MCR



While MCR can provide excellent results, it is currently more of an art than a science.

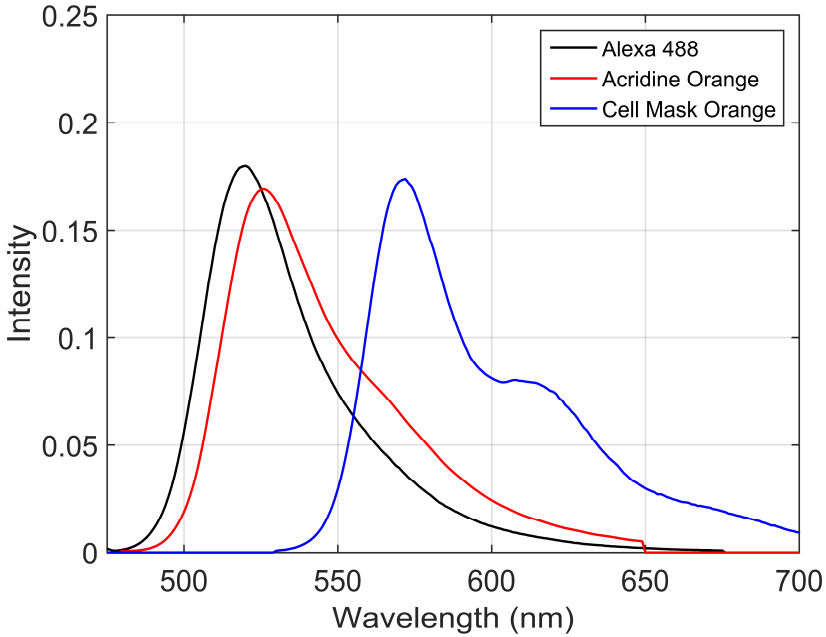
A skilled user will obtain better spectra from MCR, particularly for weaker spectral components.

Relative amounts:

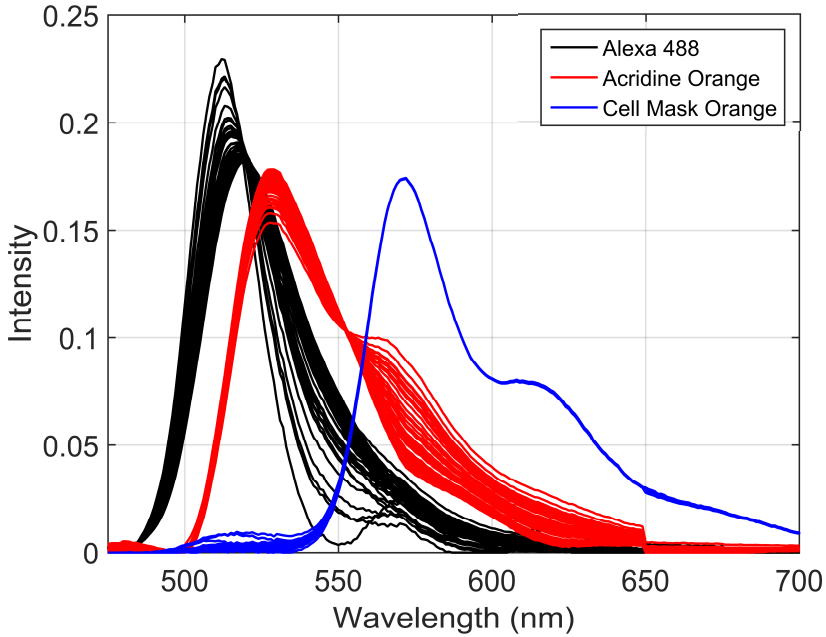
Chlorophyll B:	3.7
Chlorophyll A:	2.2
Phycobilosomes:	1.0

Multitude of Results

Pure Spectra



MCR Spectra (100 different runs)



MCR results for 100 runs on a simulated data set for a 100 x 100 pixel hyperspectral image averaging ~55000 counts for each spectrum initialized with random spectra.

Why multiple results? Two possibilities:

- Failure to converge – trapped in a local minimum
- Rotational ambiguity – results mathematically equally good

Rotational Ambiguity

Sandia National Laboratories

Also known as the rotation problem:

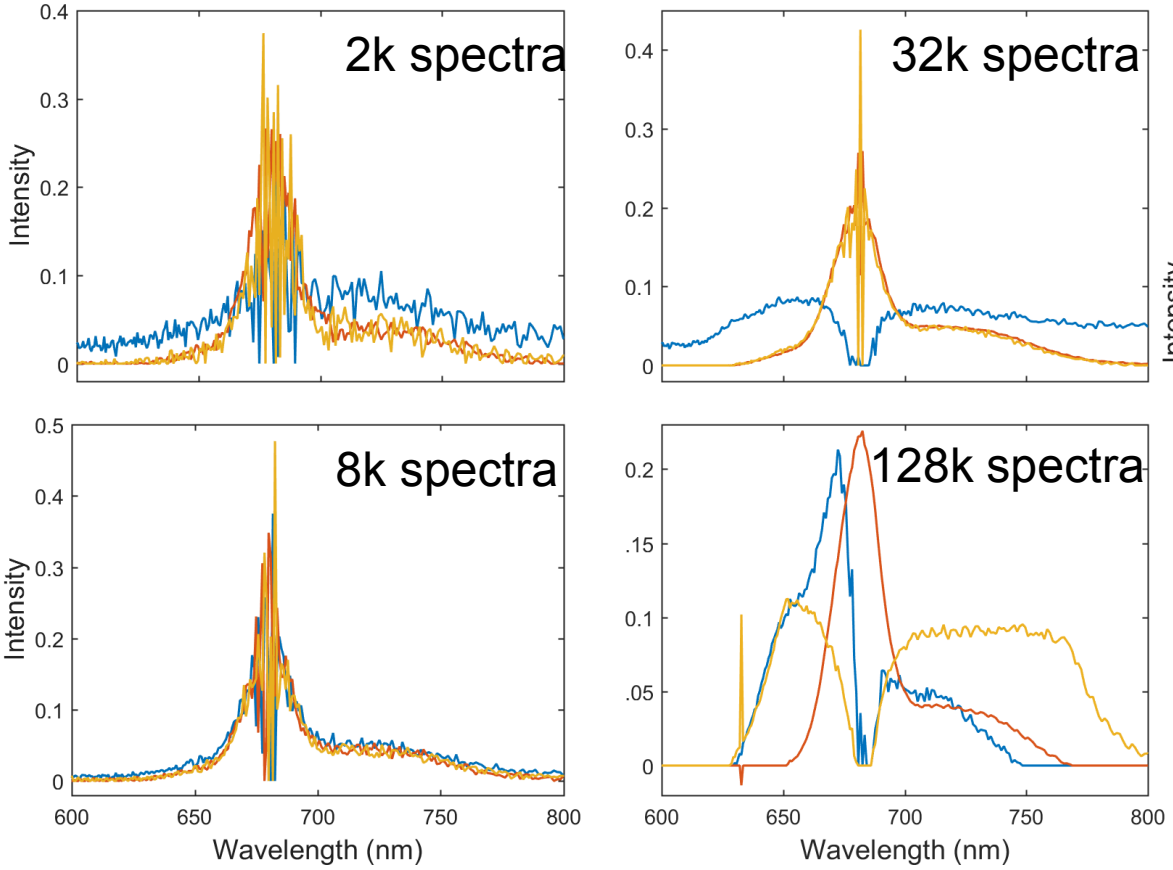
- Construct an invertible transformation matrix M_i to operate on the matrices C and S.

$$D = CS^T = (CM_i^{-1})\,(M_iS^T)$$

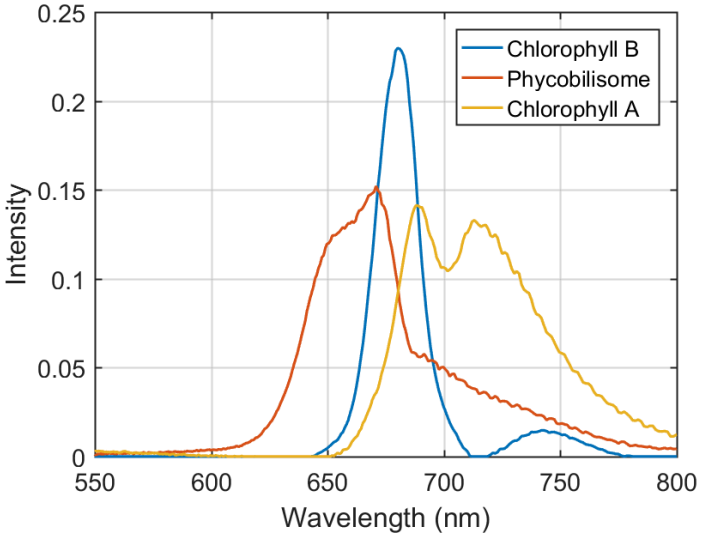
- The resulting matrices (CM_i^{-1} and M_iS^T) are an alternate solution with identical residuals

Constraints may reduce or eliminate rotational ambiguity.

Data Size



Desired Results
653k spectra

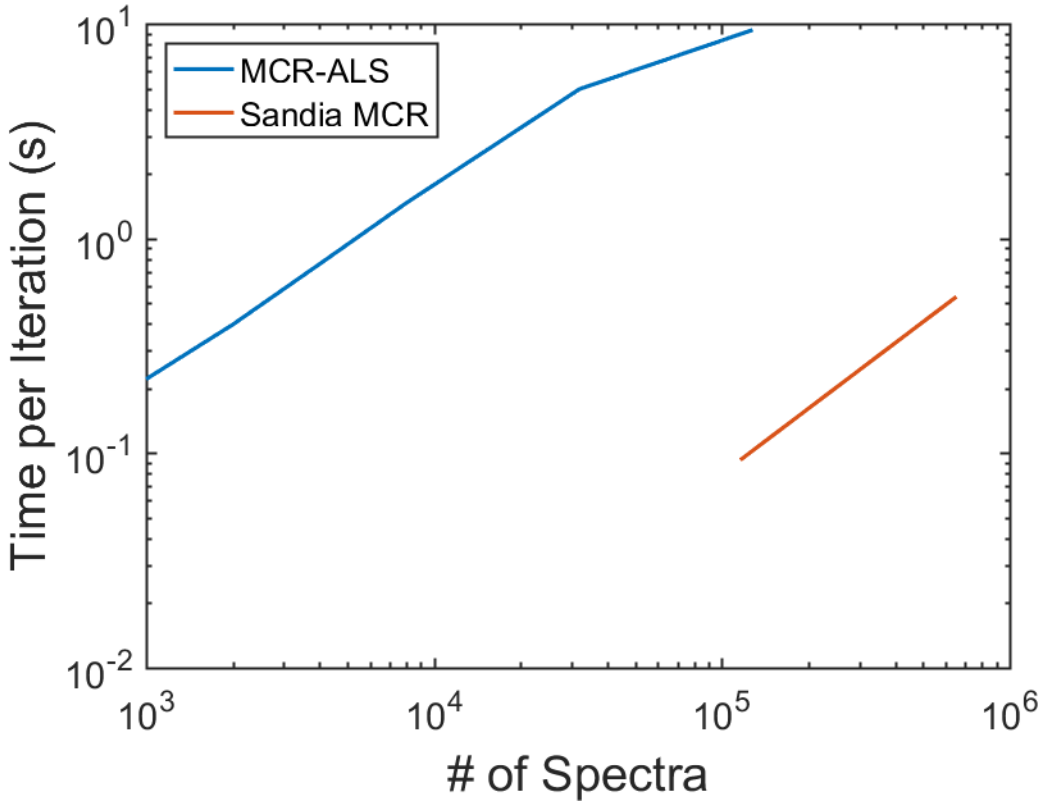


Relative amounts:

Chlorophyll B:	3.7
Chlorophyll A:	2.2
Phycobilosomes:	1.0

Working with more data provides better results

Computational Requirements



Sandia's MCR algorithm runs ~50 times faster than competing code[1] released in 2015.

Competing code[1] is widely used – cited 46 times in 2 years.

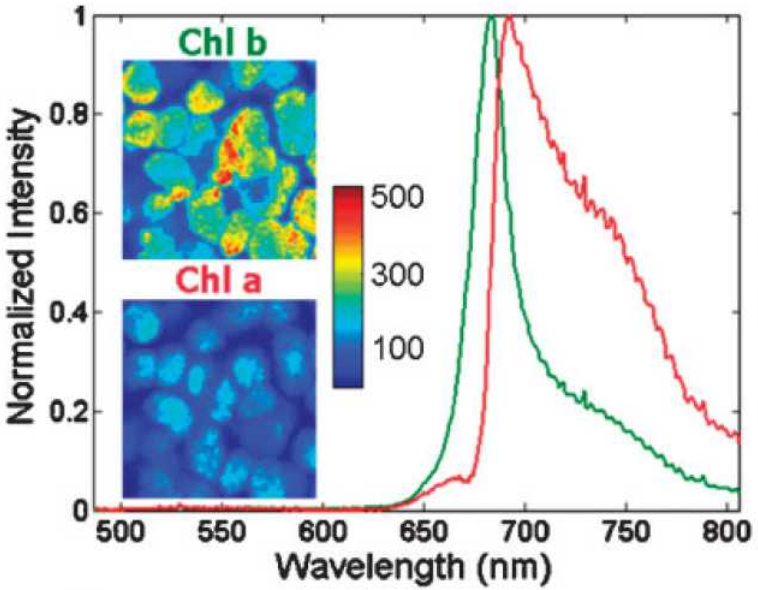
1) Jaumot, J. et al., "MCR-ALS GUI 2.0: new features and applications." *Chemometrics and Intelligent Laboratory Systems* 140 (2015): 1-12.

Computational time is on a high-end PC

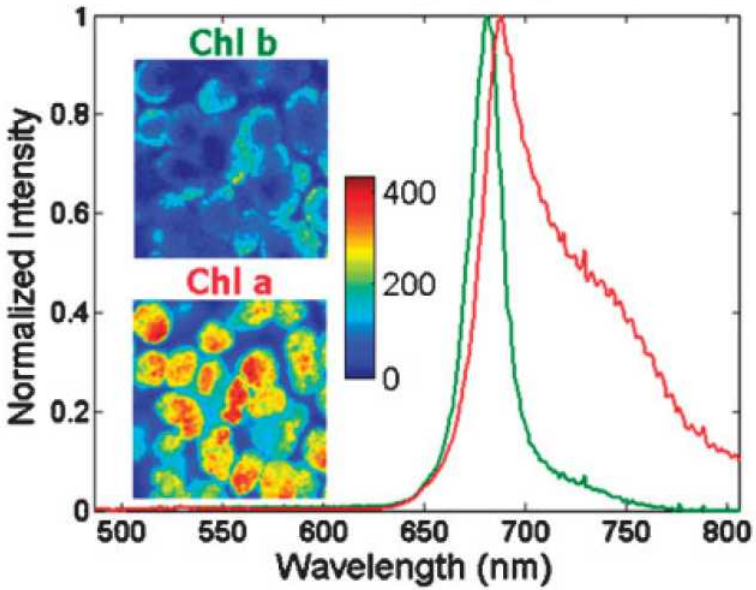
Ongoing Work – Improved Weighting

Sandia National Laboratories

Without Weighting



Weighted for Poisson & Gaussian Noise



Jones, H. D. T., et al. (2008). "Weighting hyperspectral image data for improved multivariate curve resolution results." Journal of Chemometrics **22(9-10): 482-490.**

Proper weighting makes a major difference!

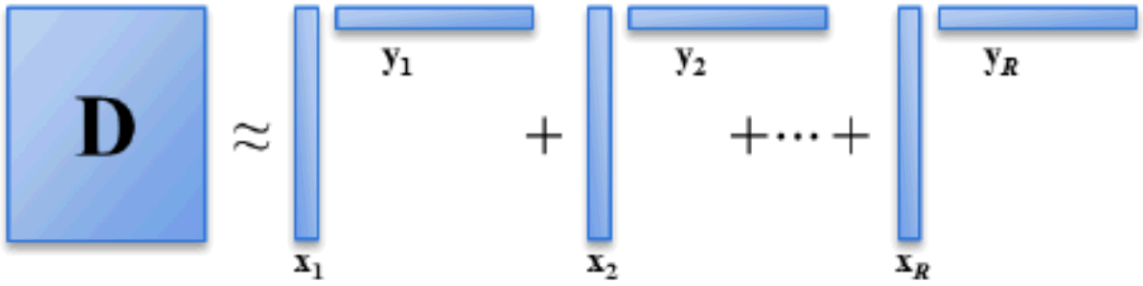
Working on improving the weighting to correctly account for all sources of noise, including the pre-processing steps.

Outline

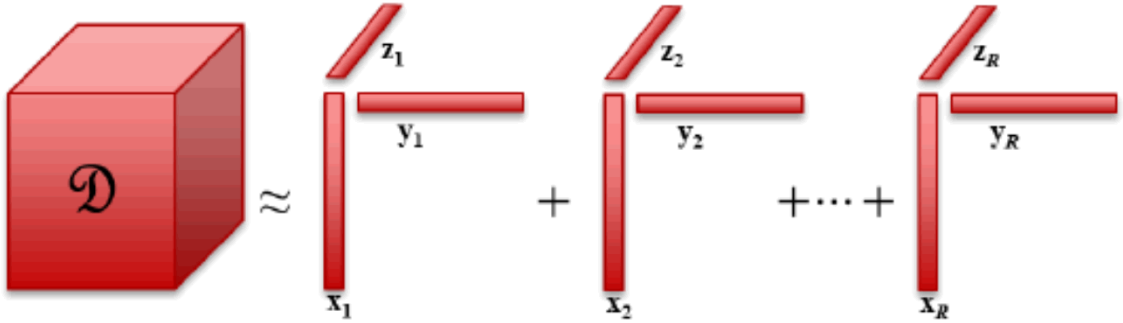
- **Why hyperspectral?** – What is hyperspectral imaging and what are its benefits?
- **Sandia's hyperspectral microscopes** – What are the systems I typically work with?
- **Analyzing hyperspectral data** – What can multivariate curve resolution (MCR) do for you?
- **Improving MCR** – Ongoing work to improve its capabilities.
- **Trilinear Data** – How to leverage additional information.

Bilinear vs. Trilinear Data

$$\mathbf{D} \approx \sum_{i=1}^{R} \mathbf{x}_i \mathbf{y}_i^{\mathrm{T}}$$



$$\mathcal{D} \approx \sum_{i=1}^{R} \mathbf{x}_i \circ \mathbf{y}_i \circ \mathbf{z}_i$$

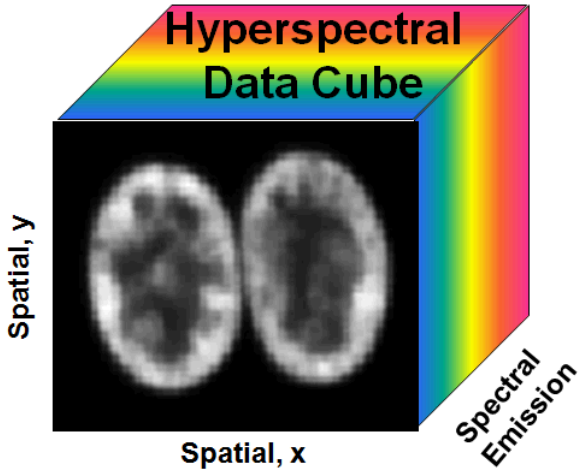


Bilinear Data – Each component in the data matrix can be expressed as the product of two vectors.

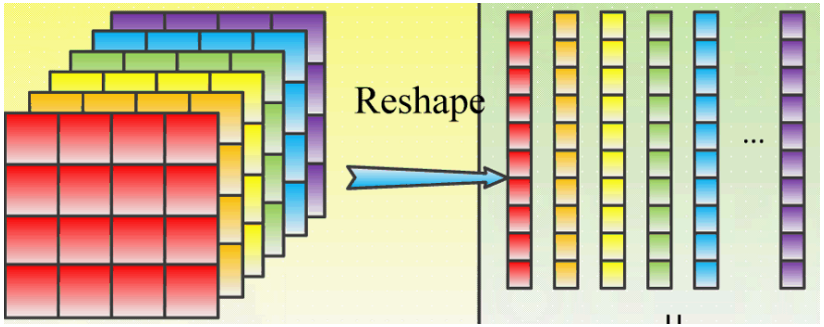
Trilinear Data – Each component in the data tensor can be expressed as the product of three vectors.

Adapted from SAND2014-1825

Hyperspectral Data Cube = Bilinear

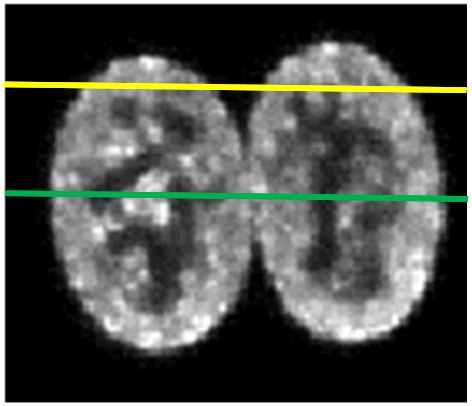


While hyperspectral data cubes are 3-dimensional, typically the data is only bilinear, not trilinear. Prior to processing with MCR, the data must be reshaped.



Adapted from Wu, Zhaojun, et al. *Journal of Electronic Imaging* 25.1 (2016): 013037-013037.

Carotenoid



In order to be trilinear, for an individual component, the cross sections at different y positions (yellow and green lines) would need to be identical.

Examples of Trilinear Data

- Hyperspectral fluorescence lifetime
  - Spatial position
  - Wavelength
  - Lifetime
- Gas Chromatography
  - Elution Time
  - Mass spectrum
  - Multiple possibilities (sample number, run temperature)

Analyzing Trilinear Data

- MCR can be applied to trilinear data, but better methods exist
  - Trilinear data can always be reshaped to generate a bilinear data set.
  - Doing so forfeits one of the great advantages of trilinear data. Trilinear analysis eliminates the rotational ambiguity problem common to bilinear data.
- Trilinear methods exist
  - Parafac, Tucker3 algorithms are examples
  - For trilinear data analysis, talk to Mark van Benthem (1814)

Summary

- Hyperspectral microscopy and MCR are valuable tools
- Hyperspectral STED will provide super-resolution capability
- Improvements to MCR will:
  - Make MCR less of an art
  - Improve estimation of the weaker components

Acknowledgements

Collaborators or prior research:

Microscopy: Jeri A. Timlin, Michael B. Sinclair, Jesse S. Aaron

MCR: Mark Van Benthem, Jeri A. Timlin, Michael R. Keenan, Paul G. Kotula

Funding

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. Support is acknowledged from Sandia National Laboratories' Laboratory Directed Research and Development project "Unmasking Hidden Compounds within Hyperspectral Images"