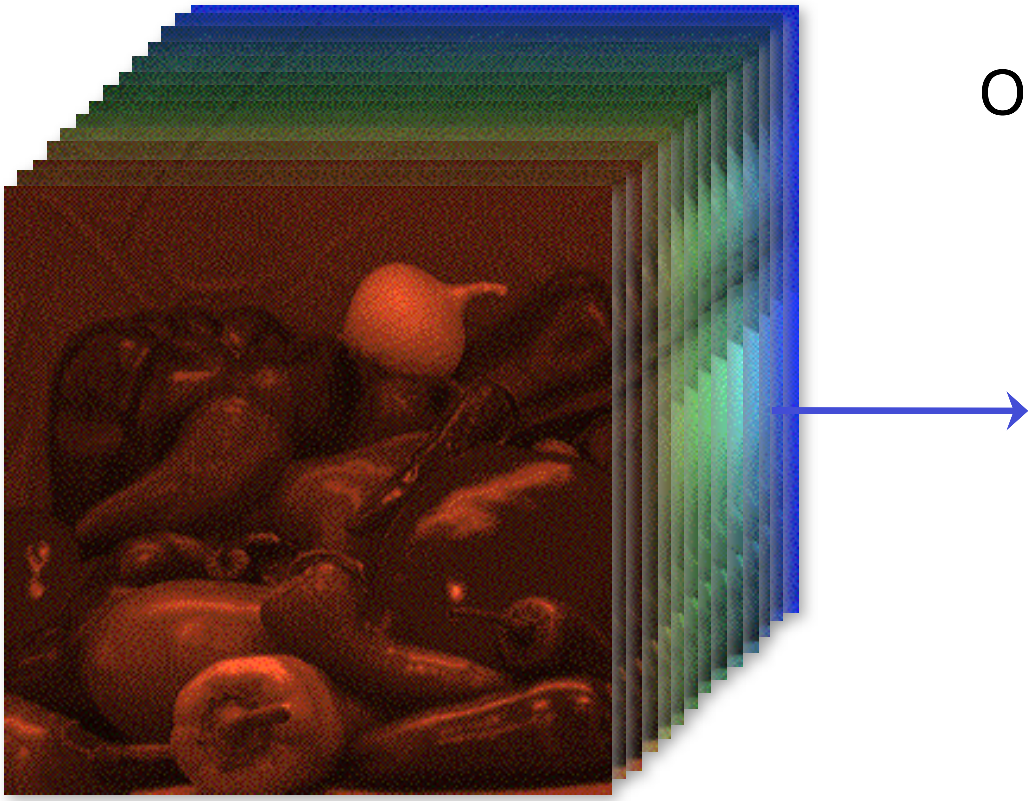
Non-Homogeneous Hidden Markov Chain Models for Wavelet-Based Hyperspectral Image Processing

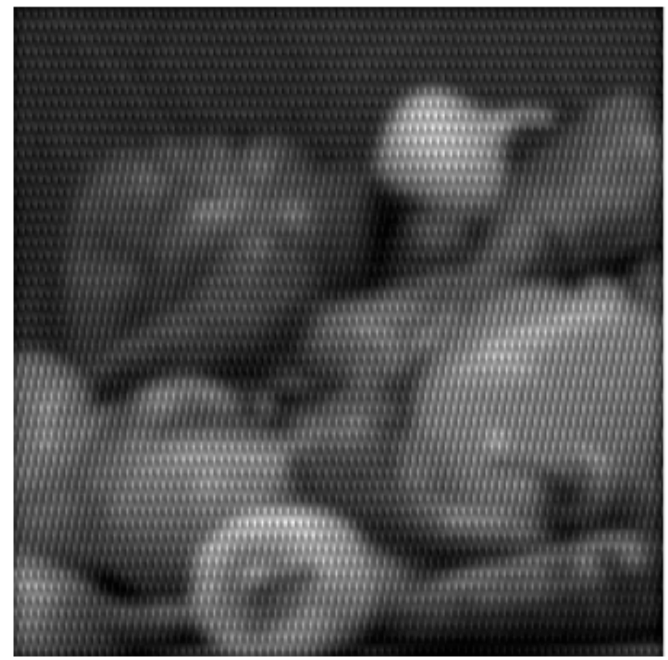
Marco F. Duarte
Mario Parente

UMASS AMHERST

Hyperspectral Imaging

One signal/image per band

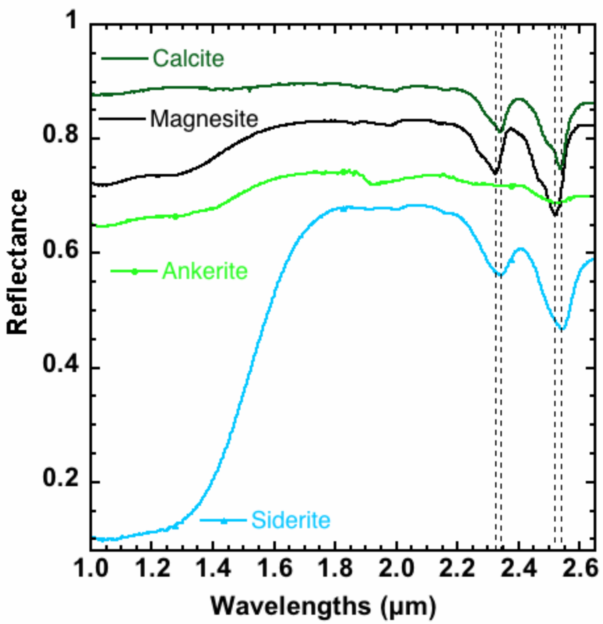




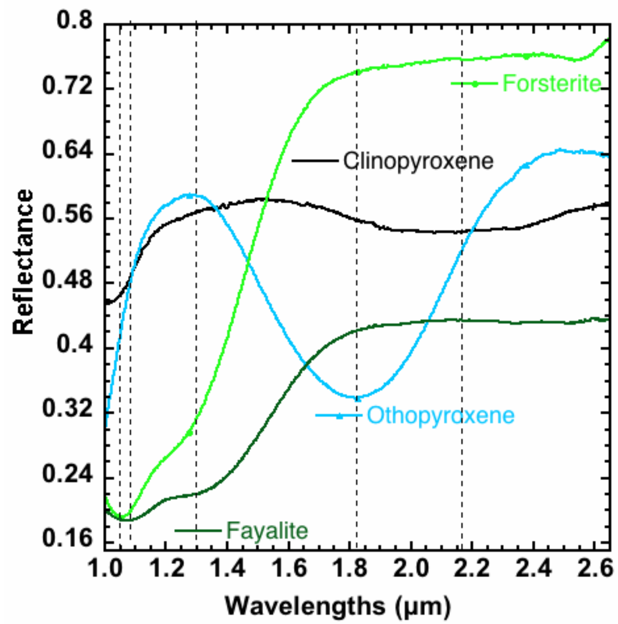
Hyperspectral datacube

Spectrum at each pixel represents
composition/physical state of subject
(remote sensing, industrial process monitoring, etc.)

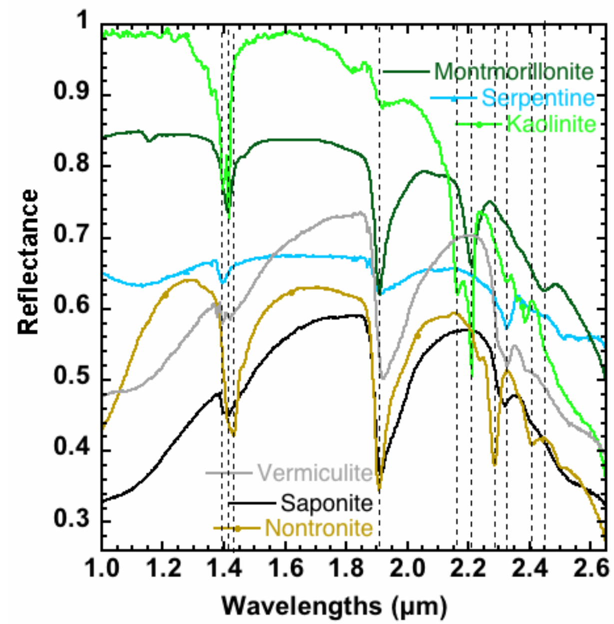
Hyperspectral Signatures



Igneous minerals



Carbonate minerals

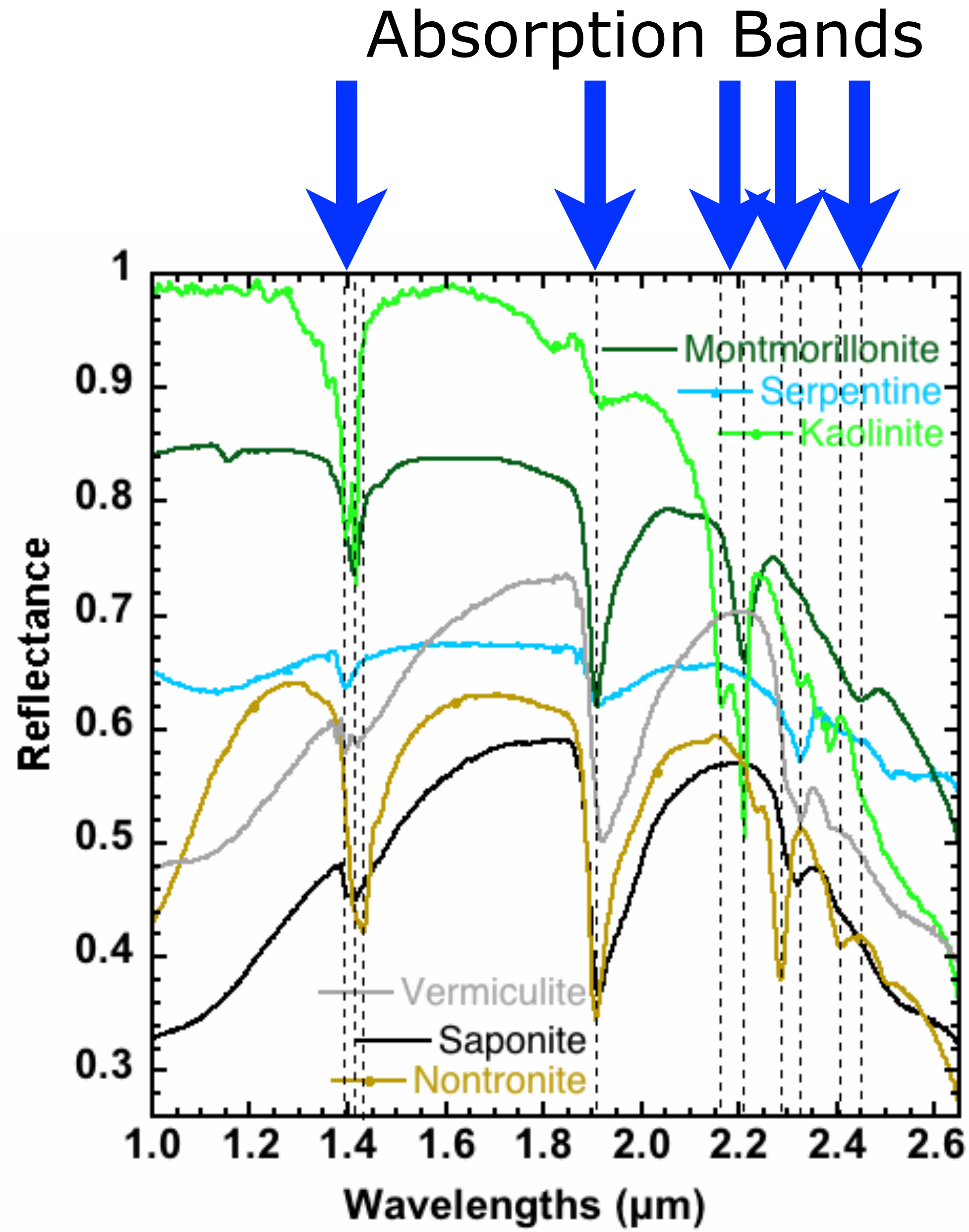


Phyllosilicate minerals (clays)

- Encode ***reflectivity of material surface*** over a variety of wavelengths of light (100+)
- Differences ***evident*** between materials/minerals of different classes; more ***subtle*** within a class
- Signature ***fluctuations*** used in ad-hoc fashion for material identification
- ***Positions*** and ***shapes*** provide identifiability

Hyperspectral Classification

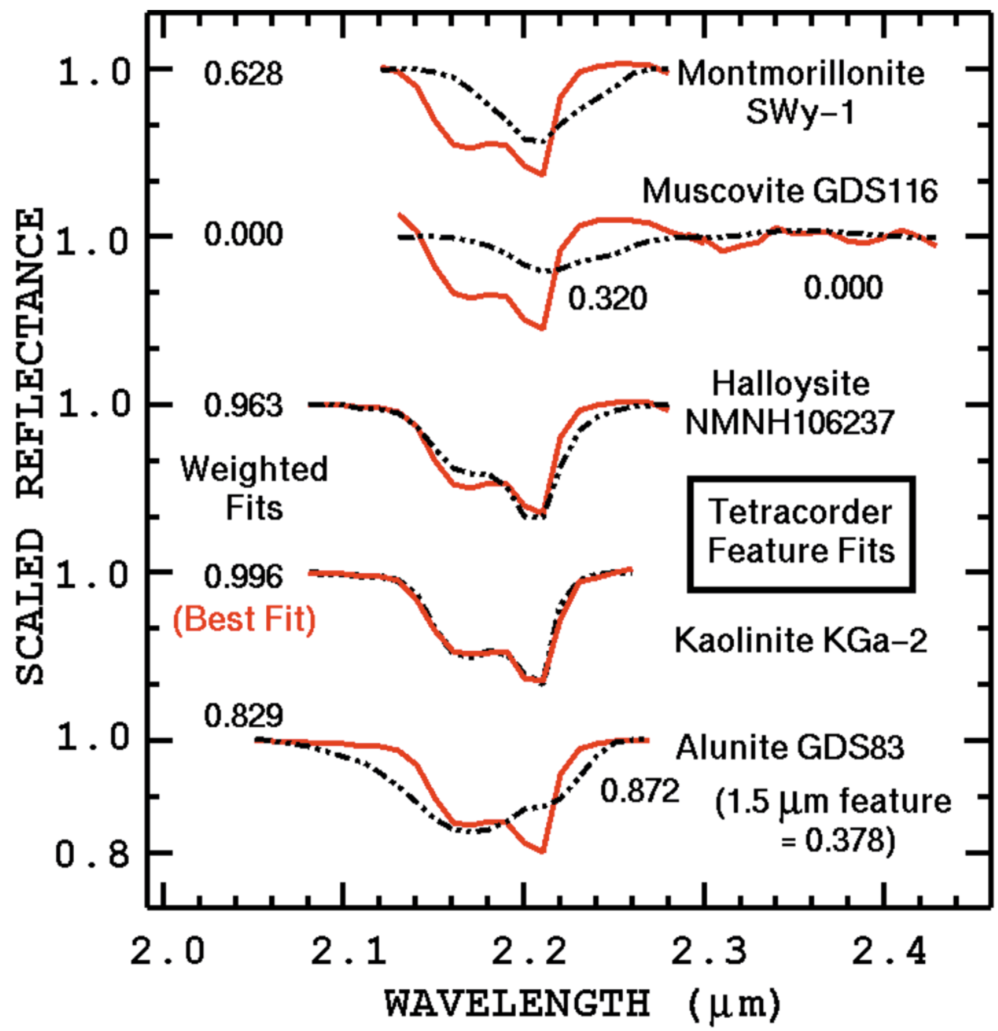
- ***Tetracorder***: List of rules to identify spectra by shape
- Rules can be arbitrarily ***complicated***
- New rules must be created for new materials
- "Difficult" cases need ***experienced analyst***



Hyperspectral Classification

- ***Tetracorder***: List of rules to identify spectra by shape
- Rules can be arbitrarily ***complicated***
- New rules must be created for new materials
- "Difficult" cases need ***experienced analyst***

[Clark et al., USGS 2003]



Hyperspectral Classification

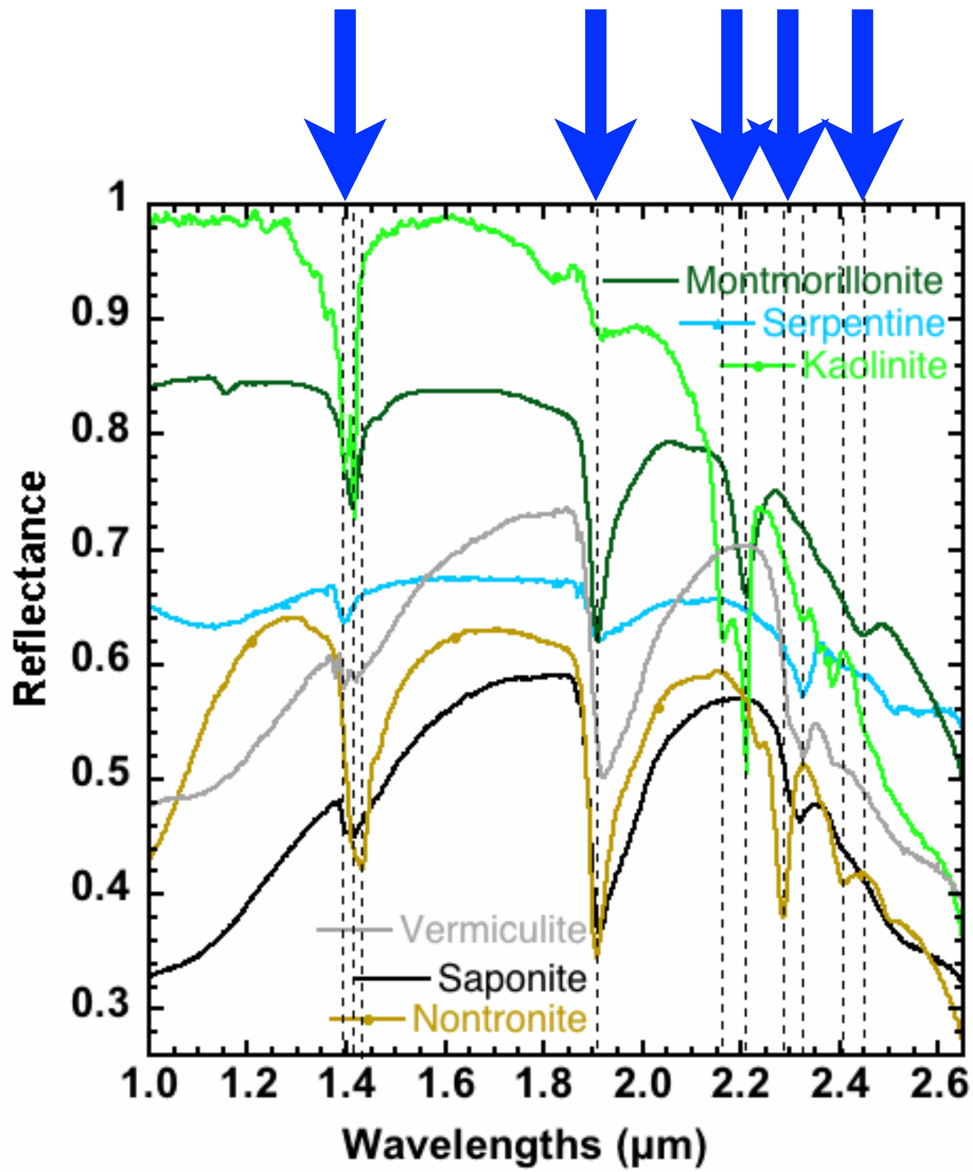
- ***Tetracorder***: List of rules to identify spectra by shape
- Rules can be arbitrarily ***complicated***
- New rules must be created for new materials
- "Difficult" cases need ***experienced analyst***

[Clark et al., USGS 2003]

group 2 # algorithm: featfit1
input library reference spectrum #=TITLE=Alunite GDS83 Na63
channels to exclude (global variable) Alunite GDS83 Na63
2 spectral features, 0 not features Dw 2.048 2.078 2.247 2.277 ct .04
continuum wavelengths, threshold (ct) Dw 1.466 1.476 1.535 1.555 ct .05
continuum wavelengths, threshold (ct) FITALL > 0.5
fit thresholds: if below 0.5, reject

Hyperspectral Classification

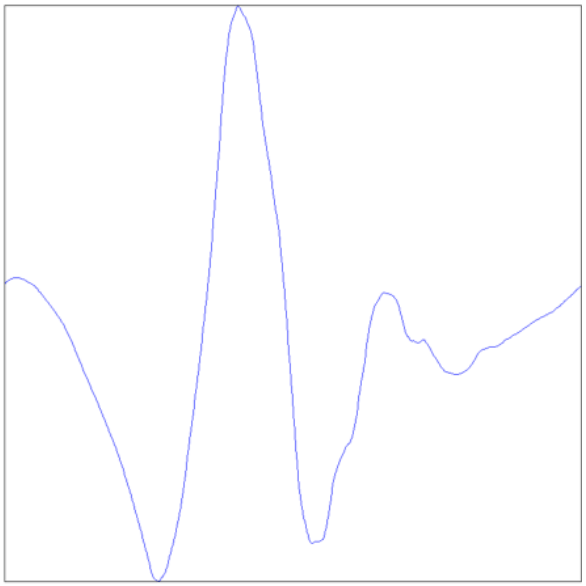
- ***Specialized*** distance metrics: spectral angle mapper, spectral divergence, etc.
  - aim to match shapes
  - ***sensitive*** to additional variations in signal from sample to sample
- How to successfully capture fluctuations in ***punctuated, piecewise smooth*** signals?



Continuous Wavelet Transform

- Mother wavelet dilated to ***scale*** s and translated to ***offset*** u:

$$\psi_{s,u}(f) = \frac{1}{\sqrt{s}}\psi\left(\frac{f-u}{s}\right)$$

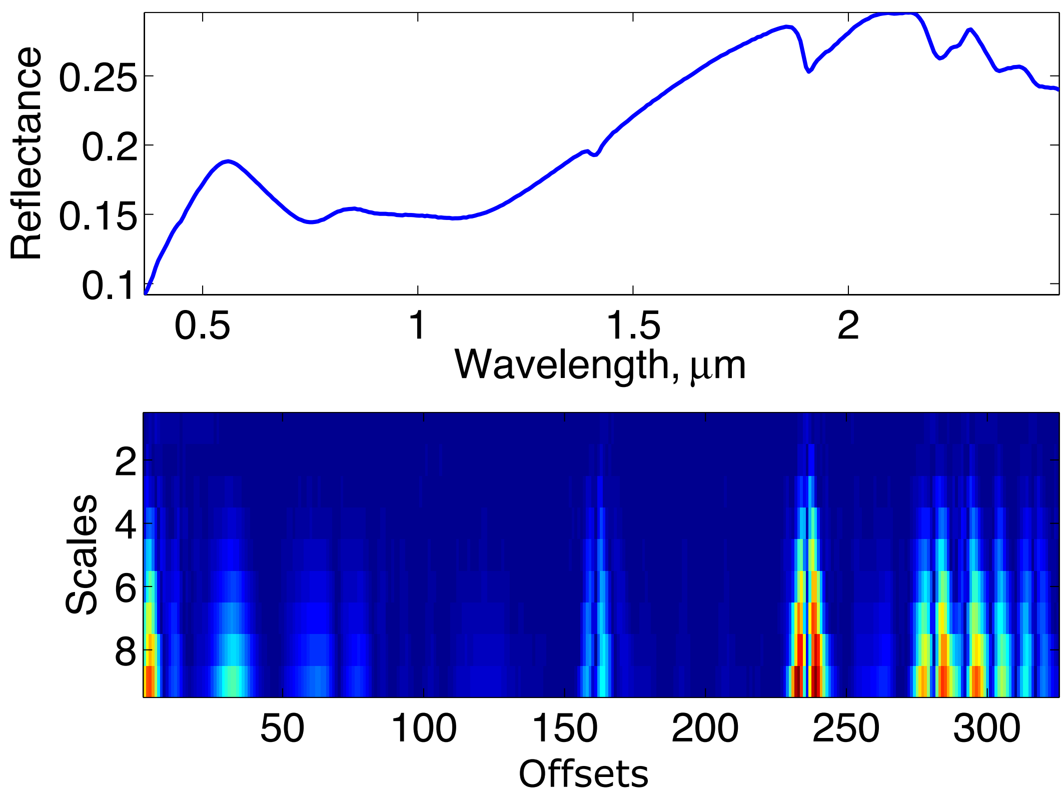


- CWT of a spectrum $x(f)$, $f \in [0, F]$, composed of wavelet coefficients $w_{s,u}$ at scales $s = 1, \ldots, S$, offsets $u = 0,\ F/N,\ 2F/N,\ \ldots,\ F\text{-}F/N$:

$$w_{s,u} = \int x(f)\psi_{s,u}(f)df$$

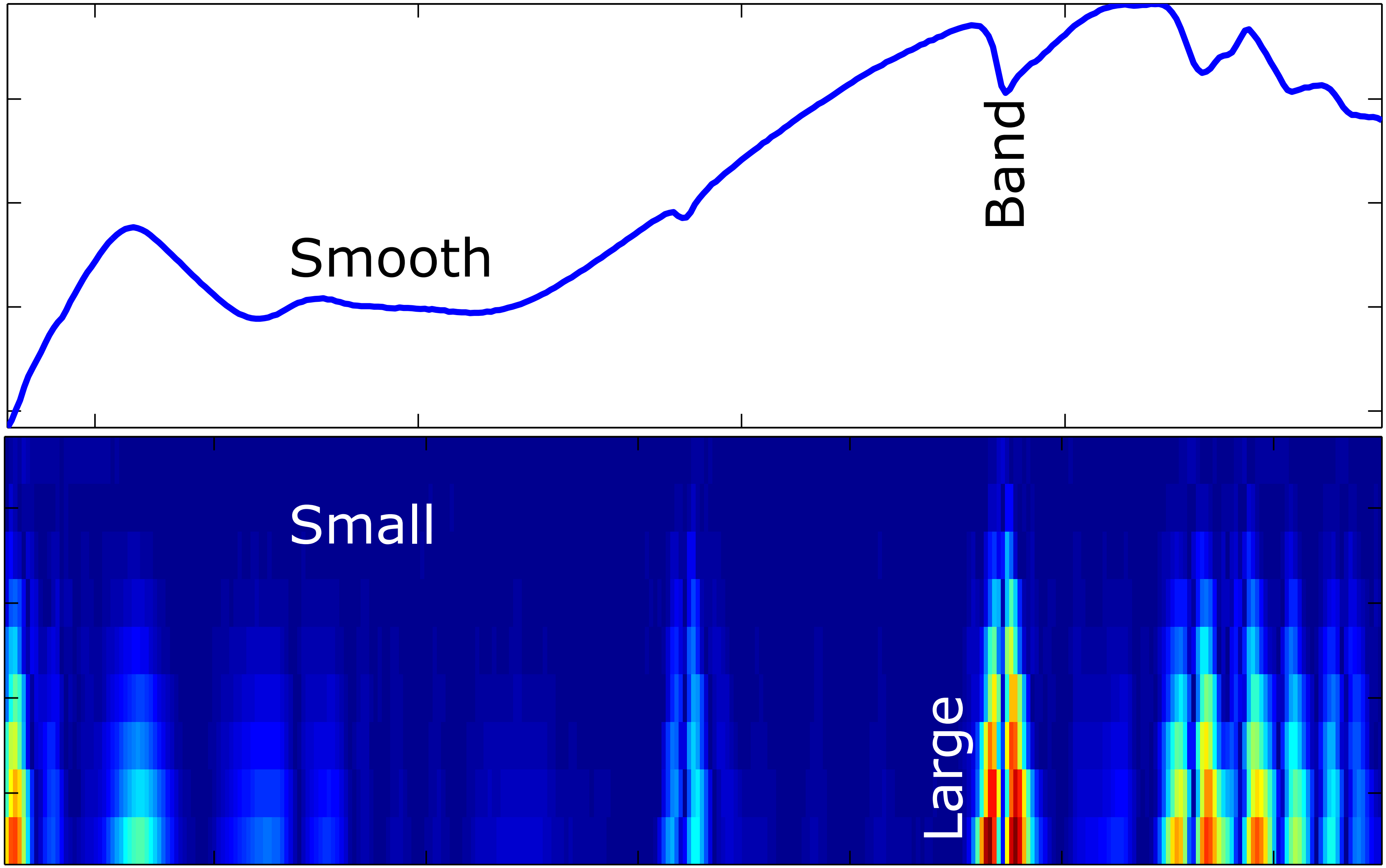
- Coefficient $w_{s,u}$ acts as a "***detector***" of fluctuations of scale s at location $f = u$

Continuous Wavelet Transform

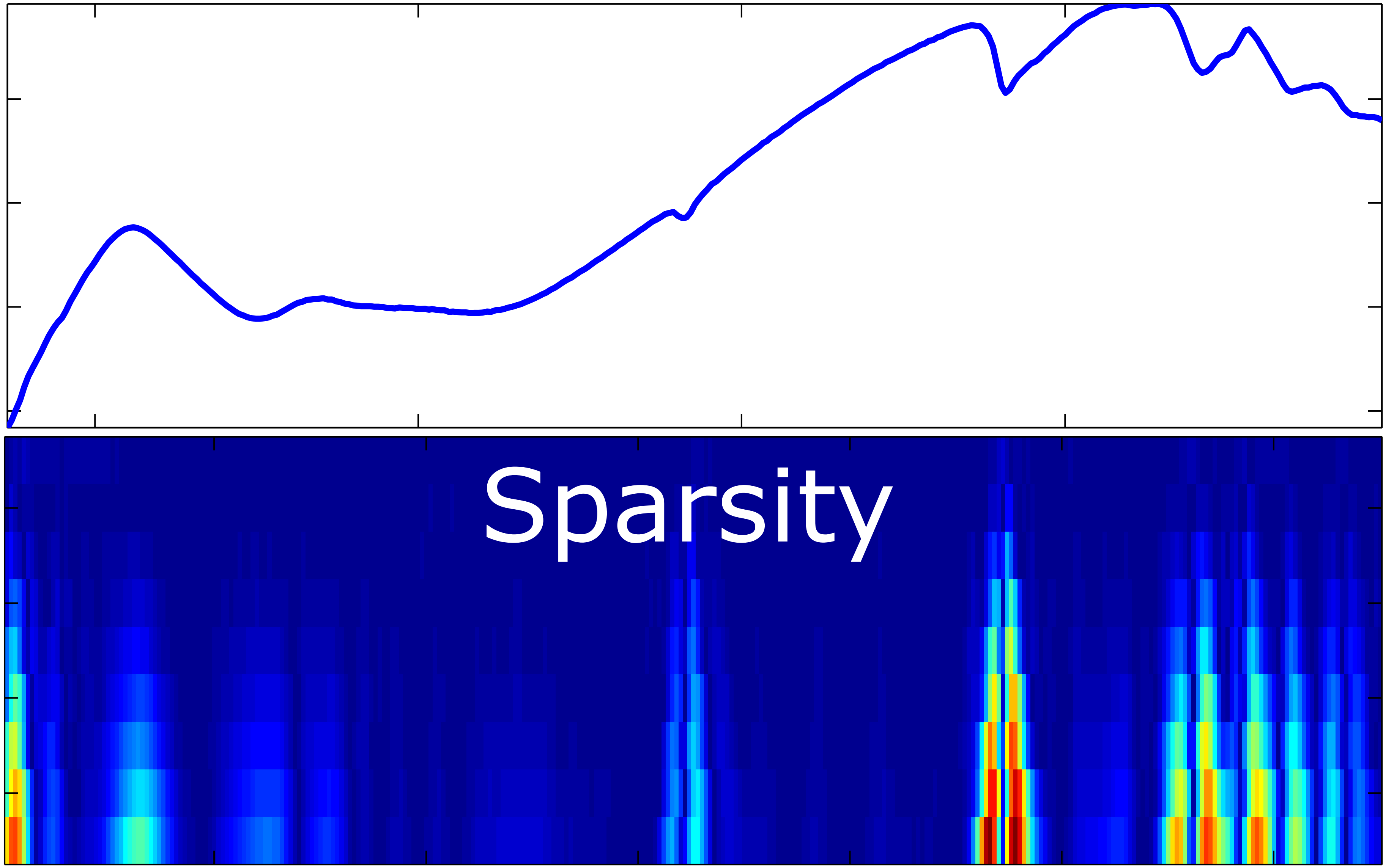


- Organize $w_{s,u}$ in a 2-D array $W \in \mathbb{R}^{S \times N}$: rows are scales, columns are offsets.
- For simplicity, offset $u = nF/N$ matched to index $n = 0,\ 1,\ \ldots,\ N$-1
- Wavelengths λ_u for indices n shown
- ***Columns*** of matrix representation give ***chains*** of parent/child wavelet coefficients

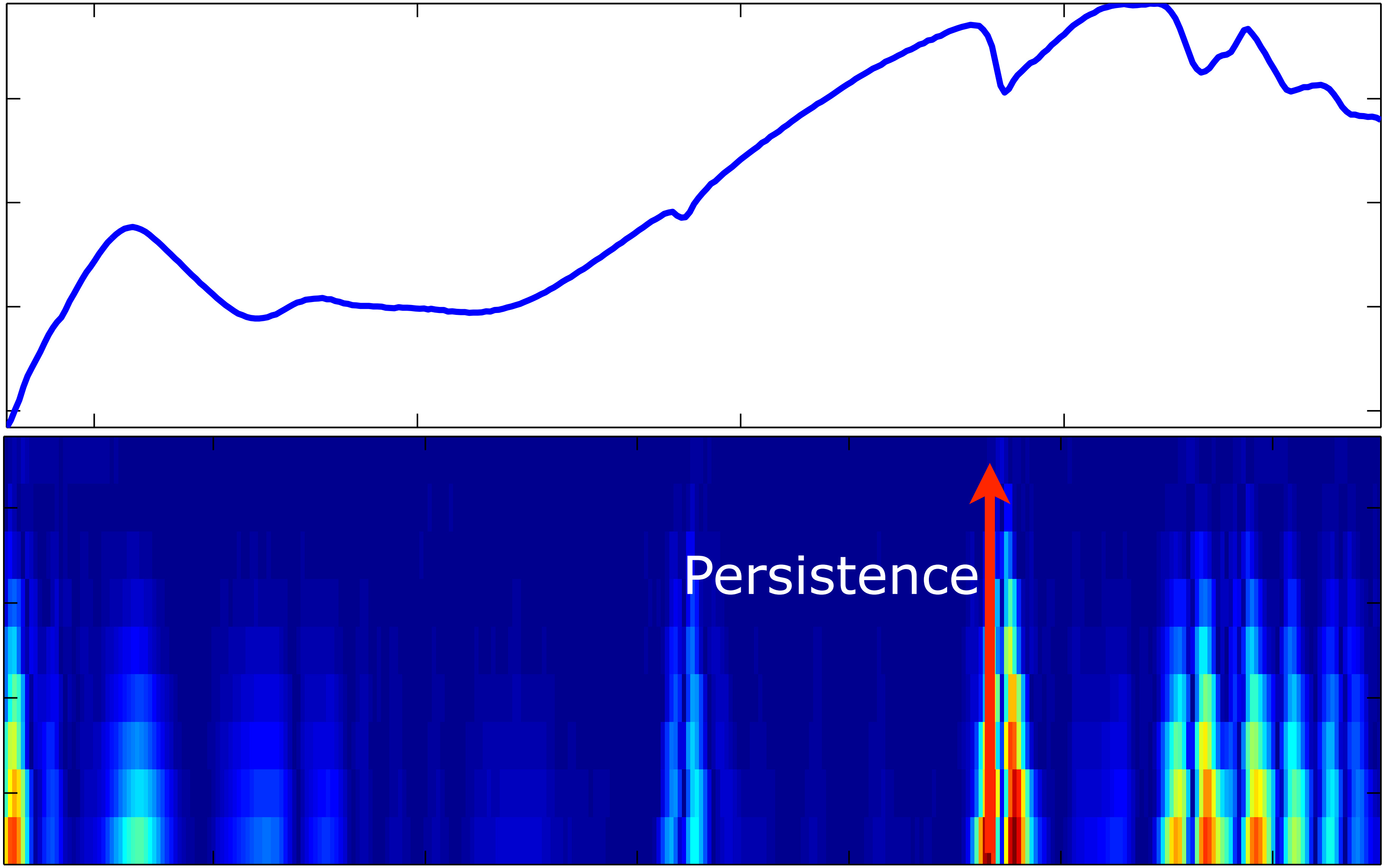
Structure of CWT Coefficients



Structure of CWT Coefficients

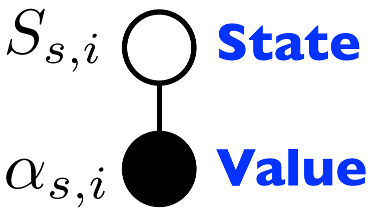


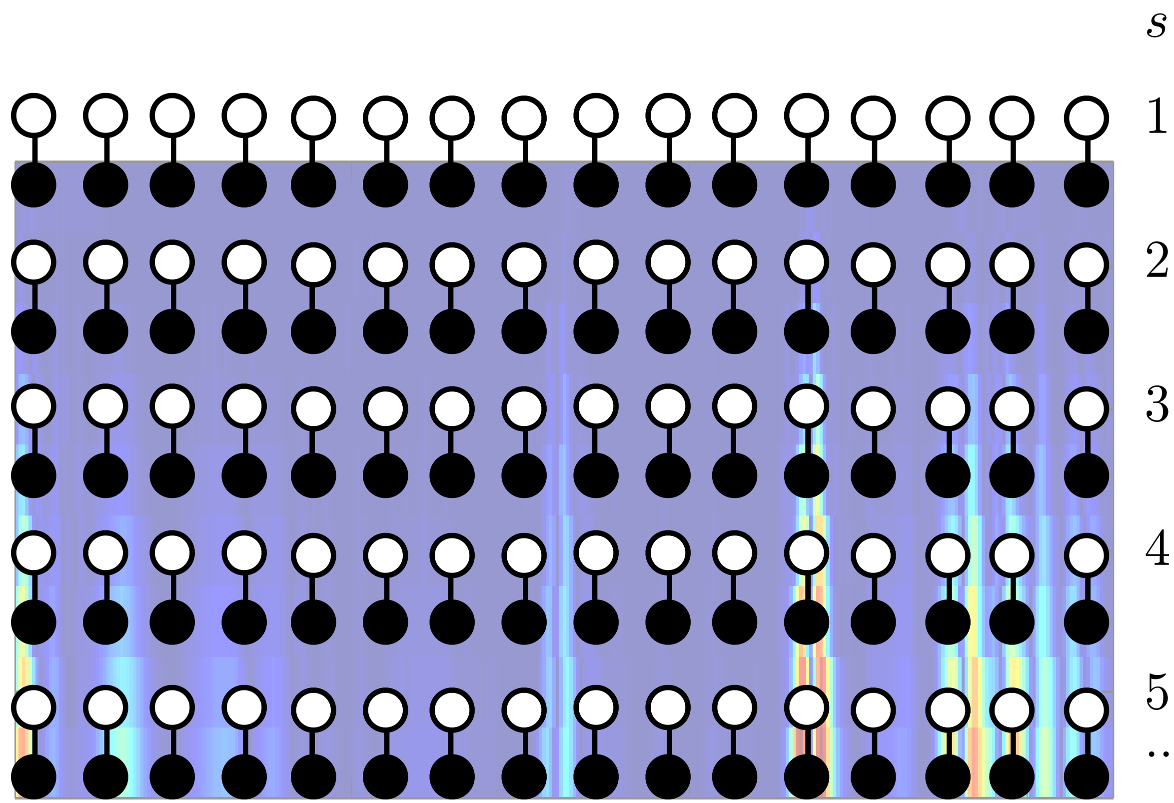
Structure of CWT Coefficients



Non-Homogeneous Hidden Markov Chains

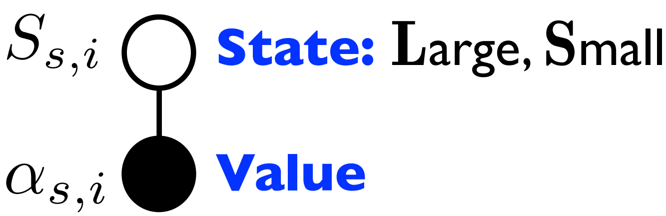
- Stochastic model to encode structure of CWT coefficients

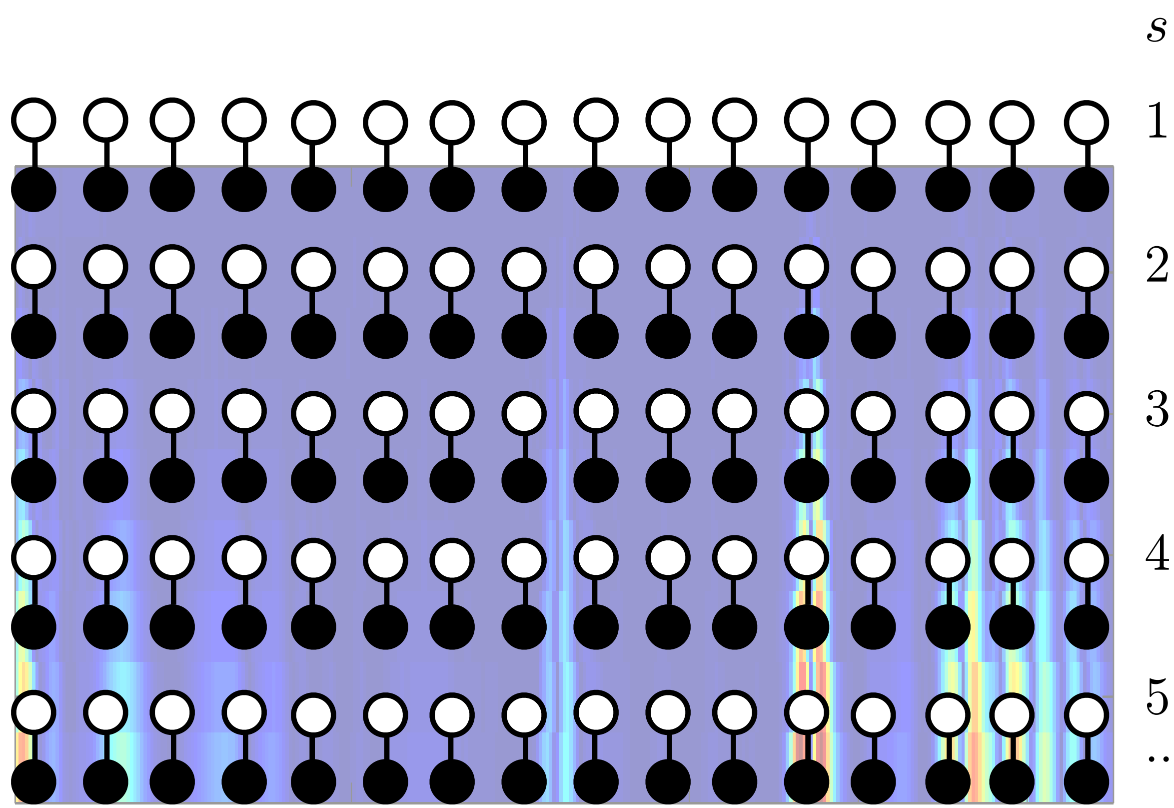




Non-Homogeneous Hidden Markov Chains

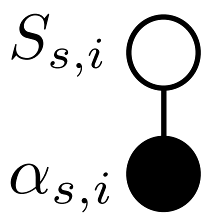
- Stochastic model to encode structure of CWT coefficients





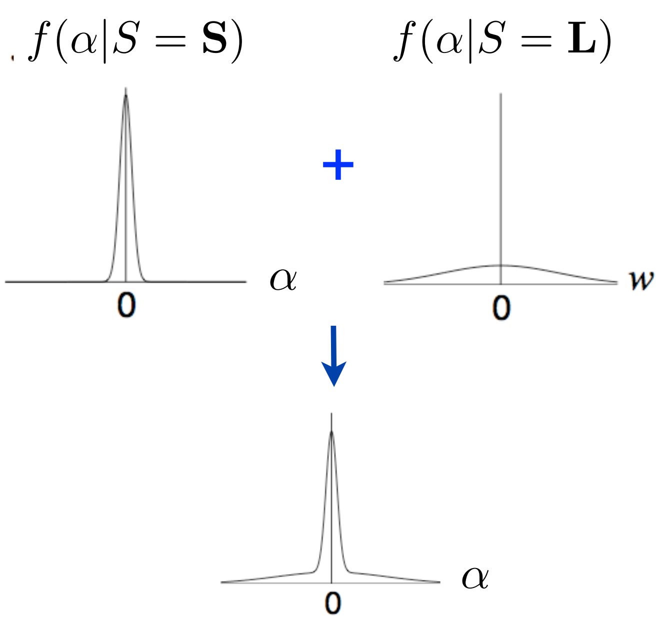
Non-Homogeneous Hidden Markov Chains

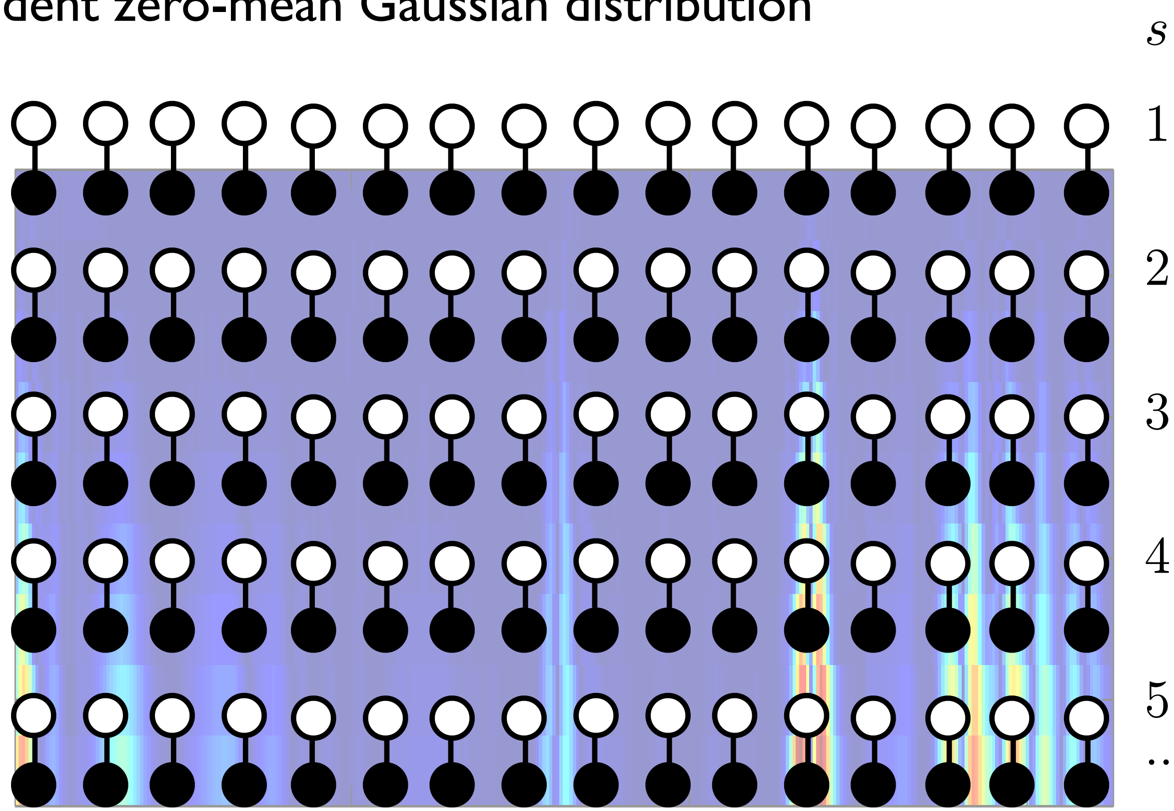
- Stochastic model to encode structure of CWT coefficients



State: Large, Small

Value: State-dependent zero-mean Gaussian distribution



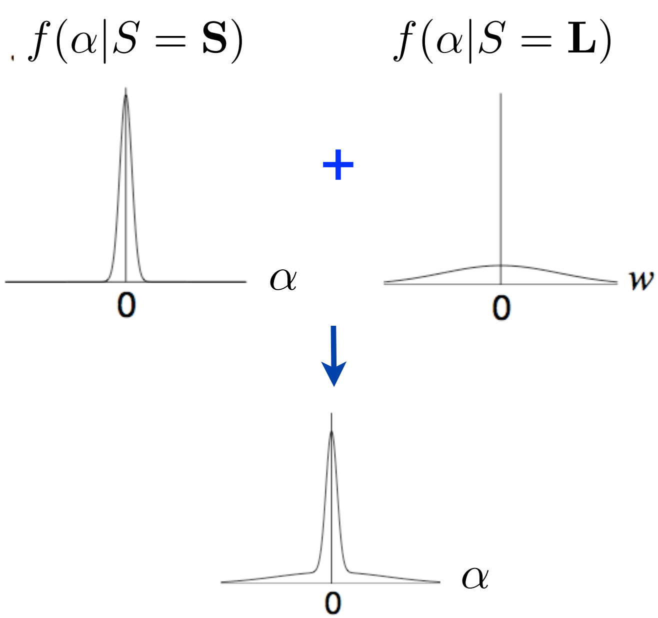


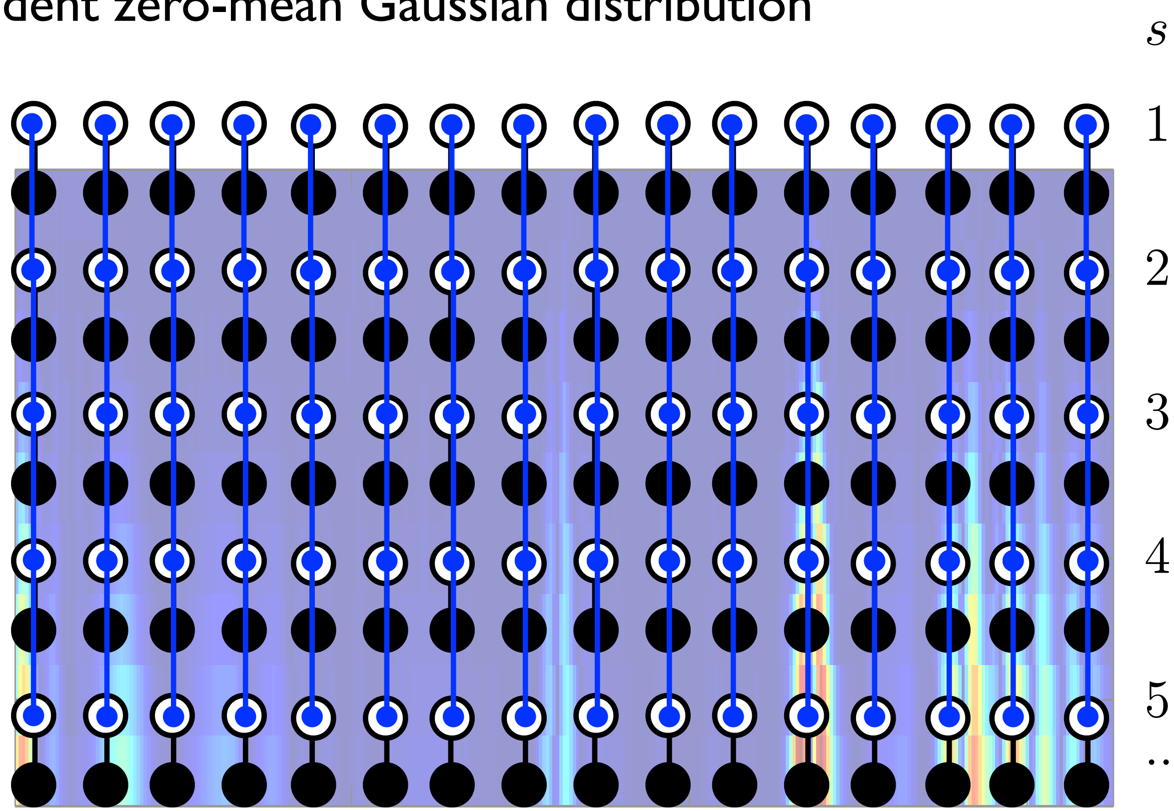
Non-Homogeneous Hidden Markov Chains

- Stochastic model to encode structure of CWT coefficients

$S_{s,i}$ **State:** **L**arge, **S**mall

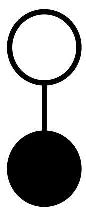
$\alpha_{s,i}$ **Value:** State-dependent zero-mean Gaussian distribution



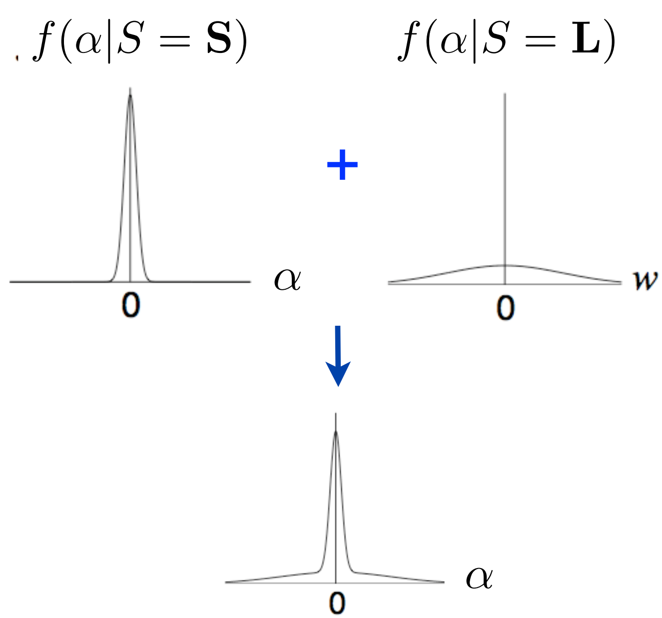


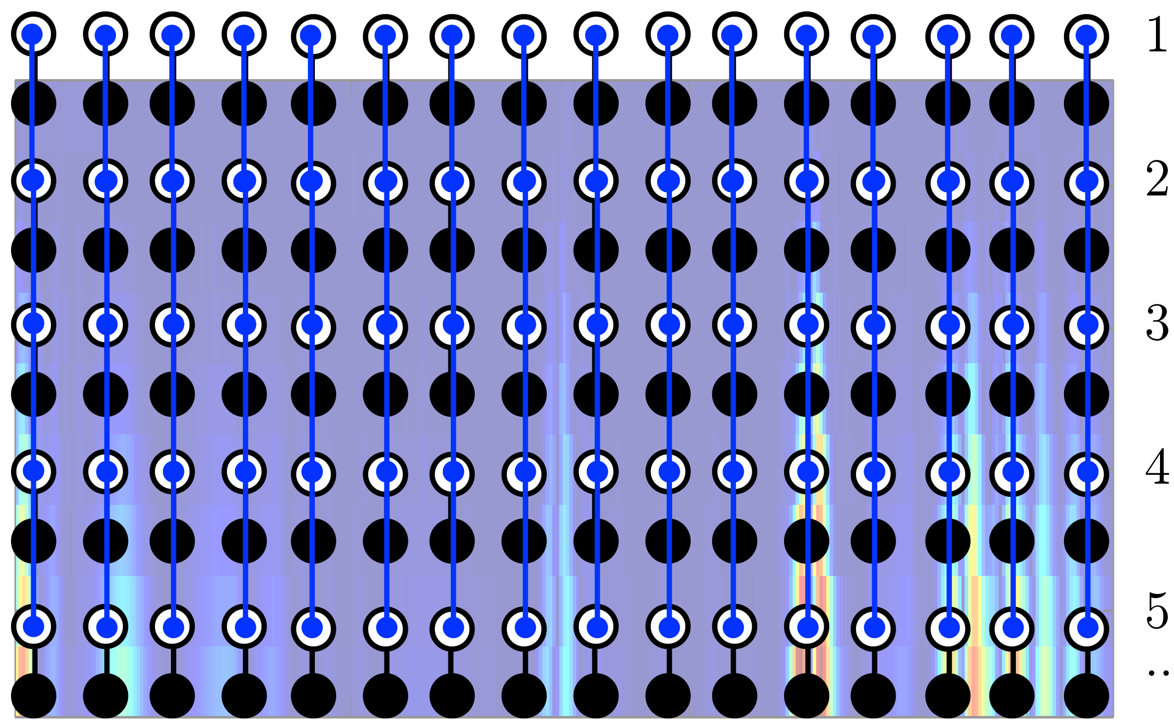
Non-Homogeneous
Hidden Markov Chains

- Stochastic model to encode structure of CWT coefficients

$S_{s,i}$  **State:** To obtain *persistence*, favor progressions $\mathbf{L} \rightarrow \mathbf{L} \rightarrow \mathbf{L} \rightarrow \ldots$ $\mathbf{S} \rightarrow \mathbf{S} \rightarrow \mathbf{S} \rightarrow \ldots$

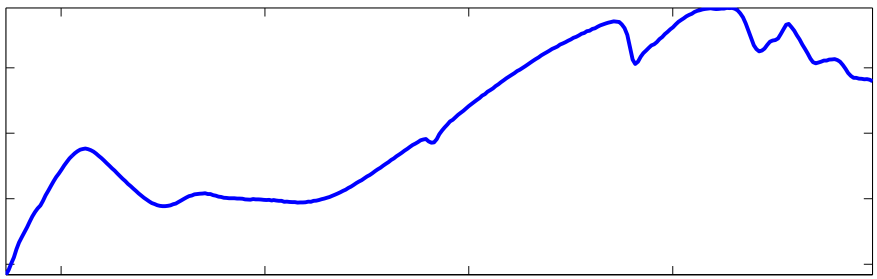
$\alpha_{s,i}$ **Value:** To obtain *decay*, reduce variances across scales

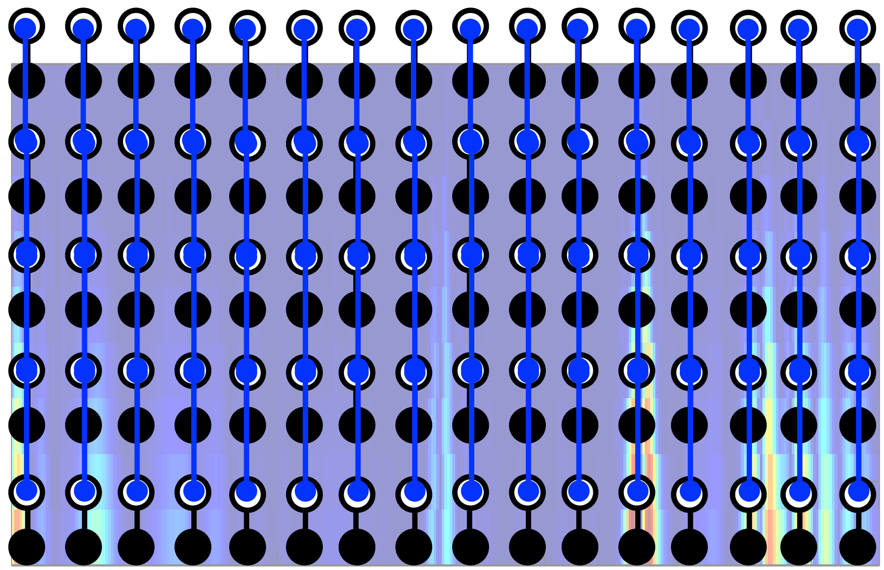




Modeling Hyperspectral Datasets

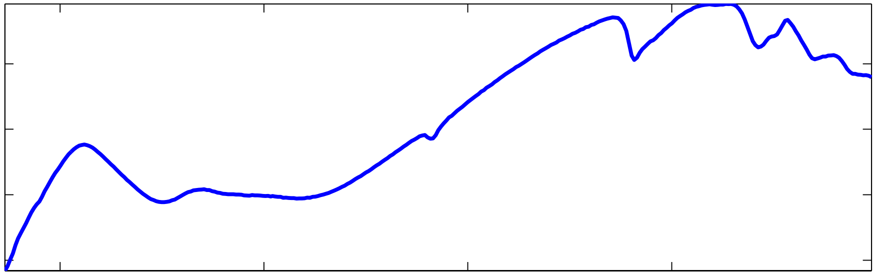
- *Why use continuous/ undecimated wavelets?* So that information at each scale is available for ***each wavelength***
- *Why separate chains for each spectra?* Because the “size” of a relevant fluctuation is ***relative to wavelength*** (e.g., absorption bands appearing in all spectra)

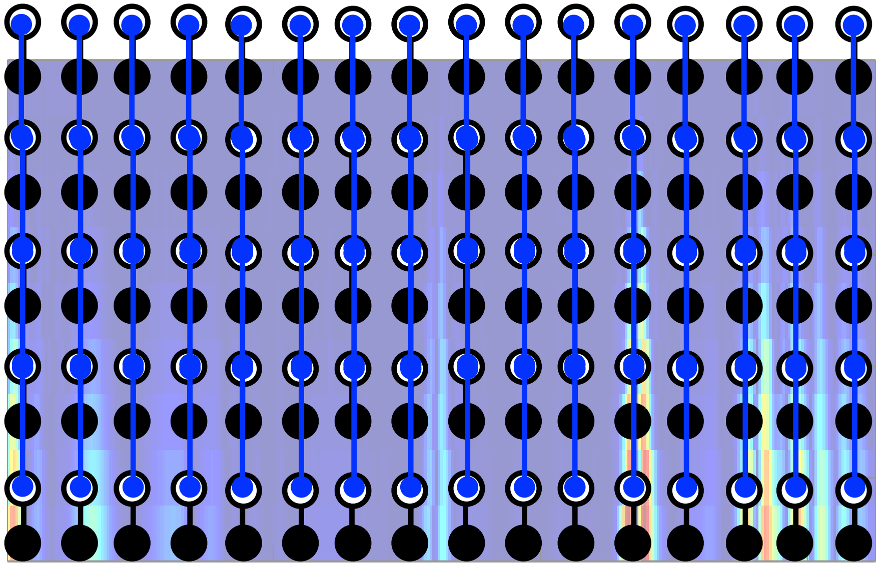




Modeling Hyperspectral Datasets

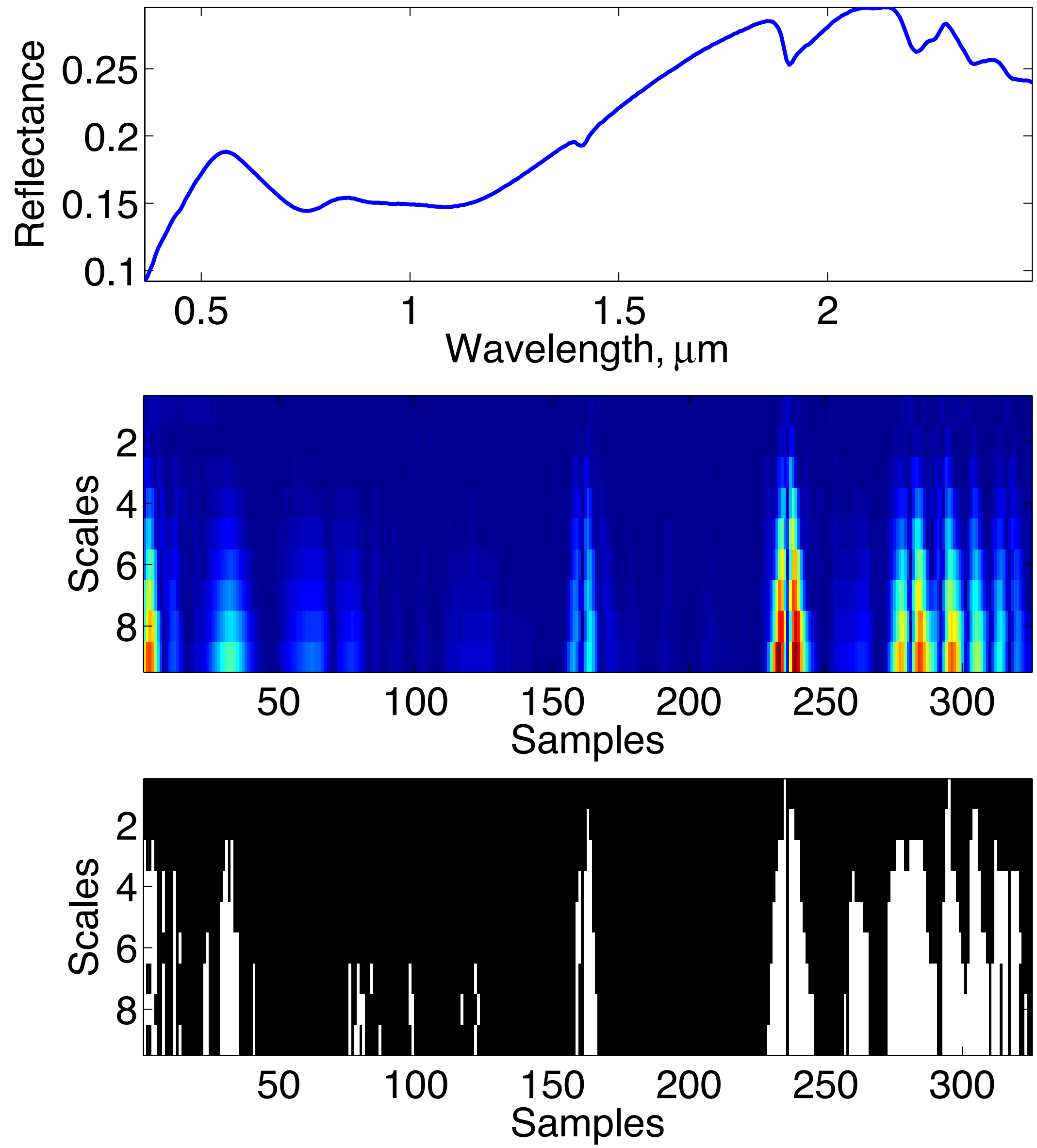
- Collect representative (***universal***) library of hyperspectral signatures (e.g. USGS for minerals)
- Extract CWT coefficients for each hyperspectral signature; collect into 2-D array
- Train an NHMC on each of the N wavelengths (array columns) over the spectral library





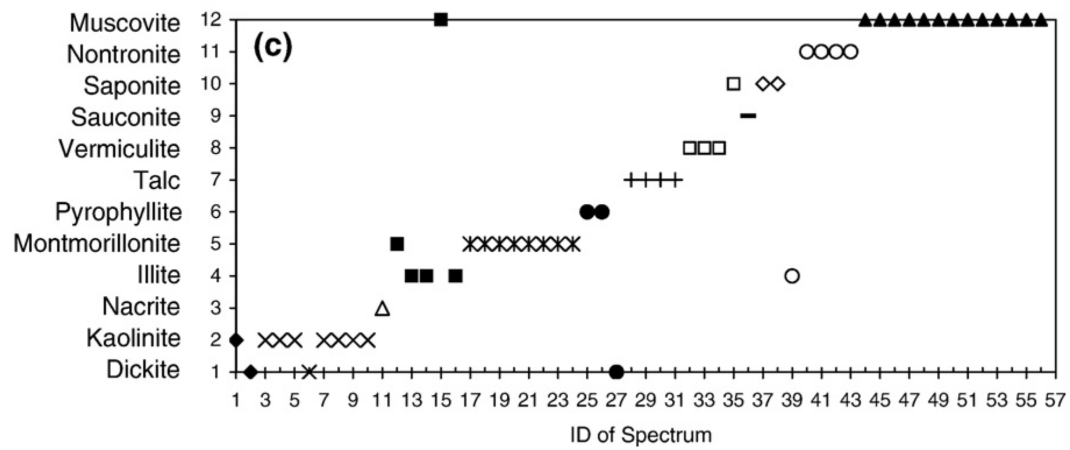
Modeling Hyperspectral Datasets

- Using learned NHMC model, generate ***state probabilities/ labels*** for each hyperspectral signature in library
- State labels provide binary information on "***interesting***" parts of the signal
- Use as ***features*** in hyperspectral signature processing (e.g., classification)

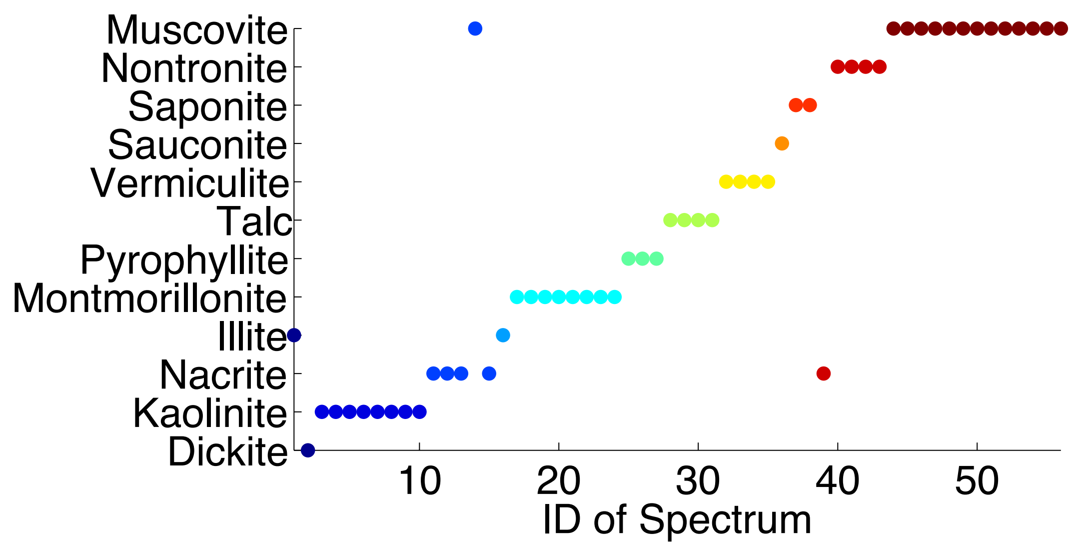


Example: Mineral Classification

- ***USGS*** spectral library with 57 clay samples from 12 classes [Rivard et al., 2008].
- One prototype/ endmember per class, classify rest by ***nearest-neighbor*** (NN) to prototypes.
- Classification ***errors*** are points that deviate from diagonal.



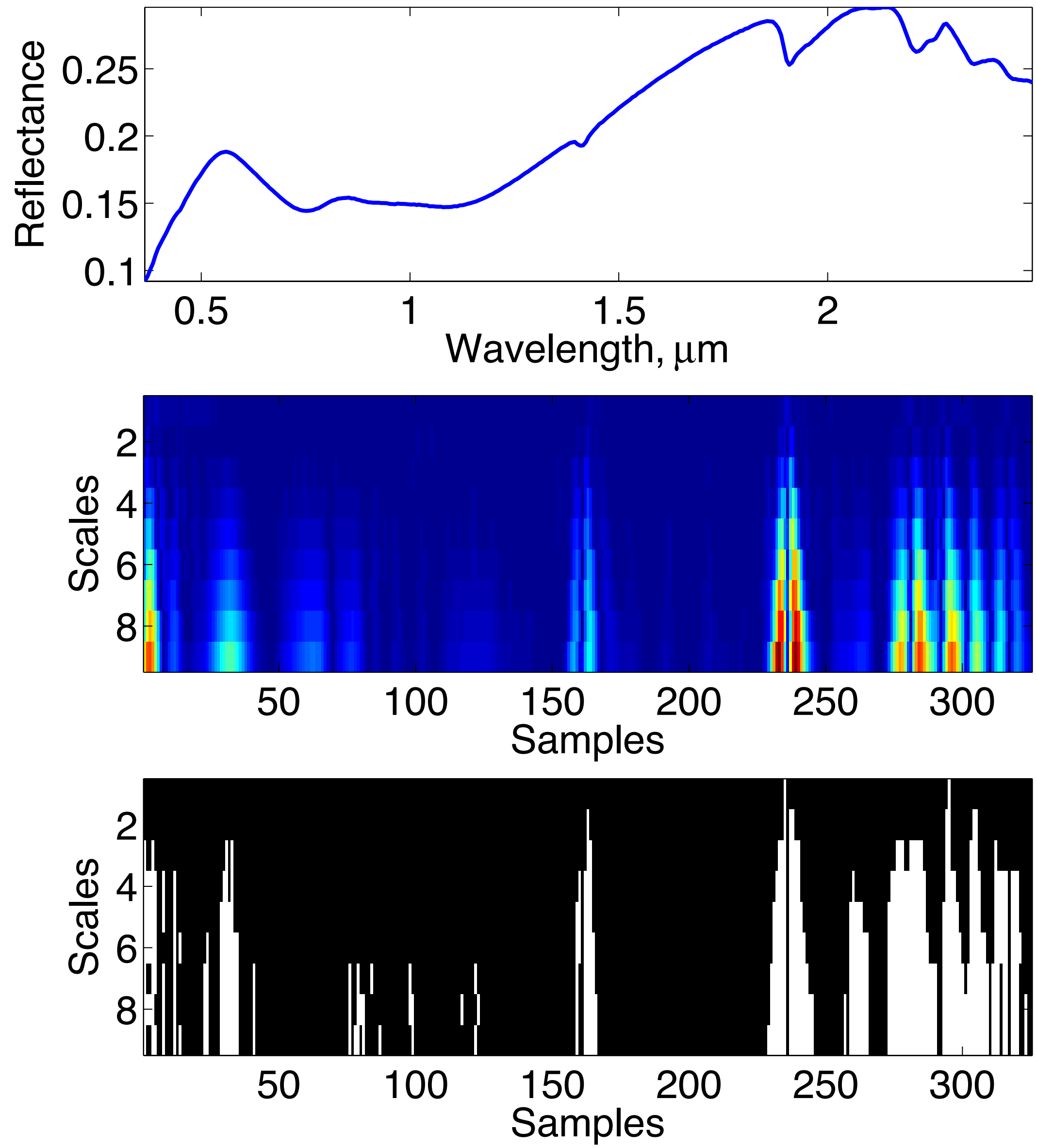
[Rivard et al., 2008] 89%



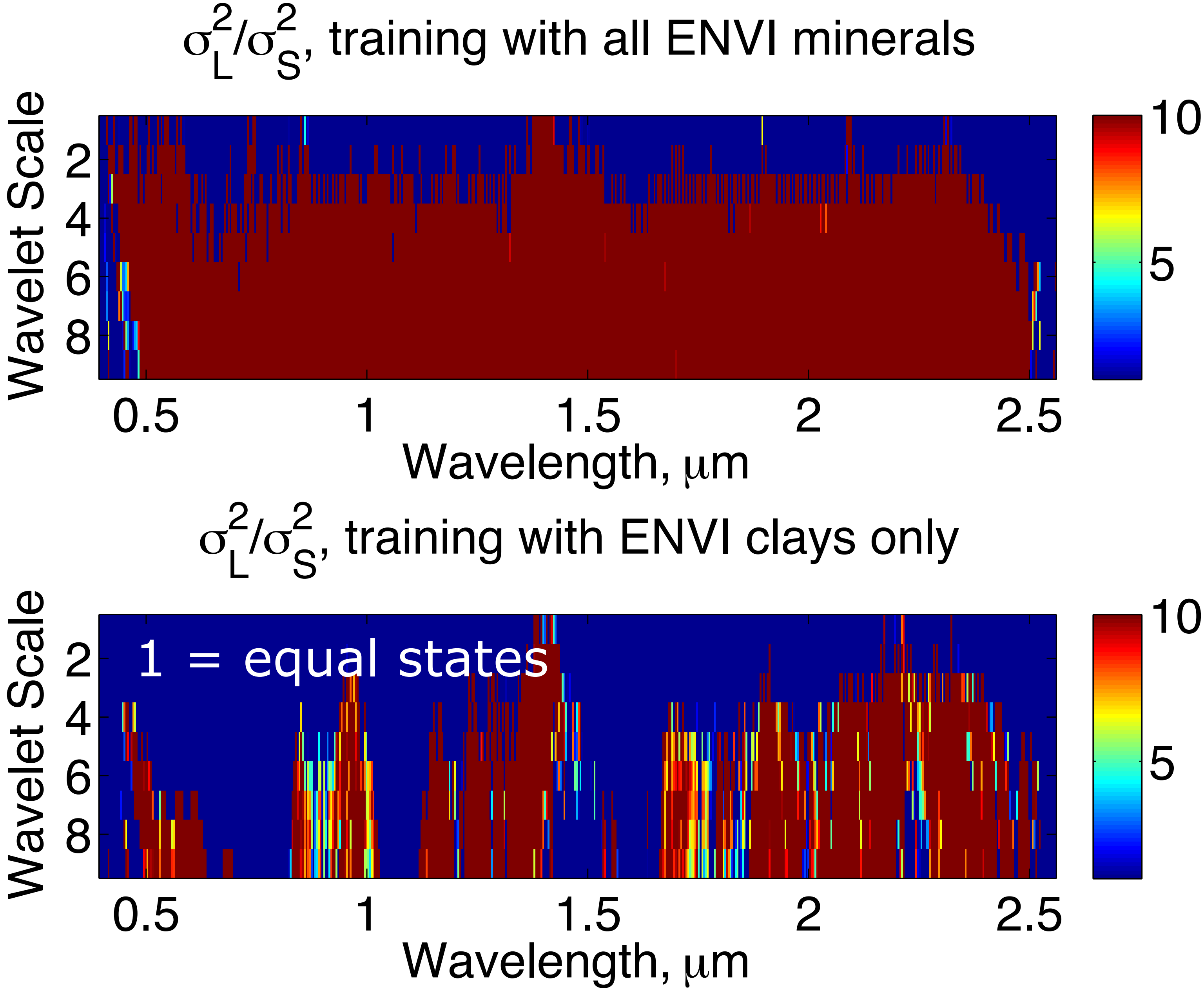
NHMC 95%

The Power of "Big Data"

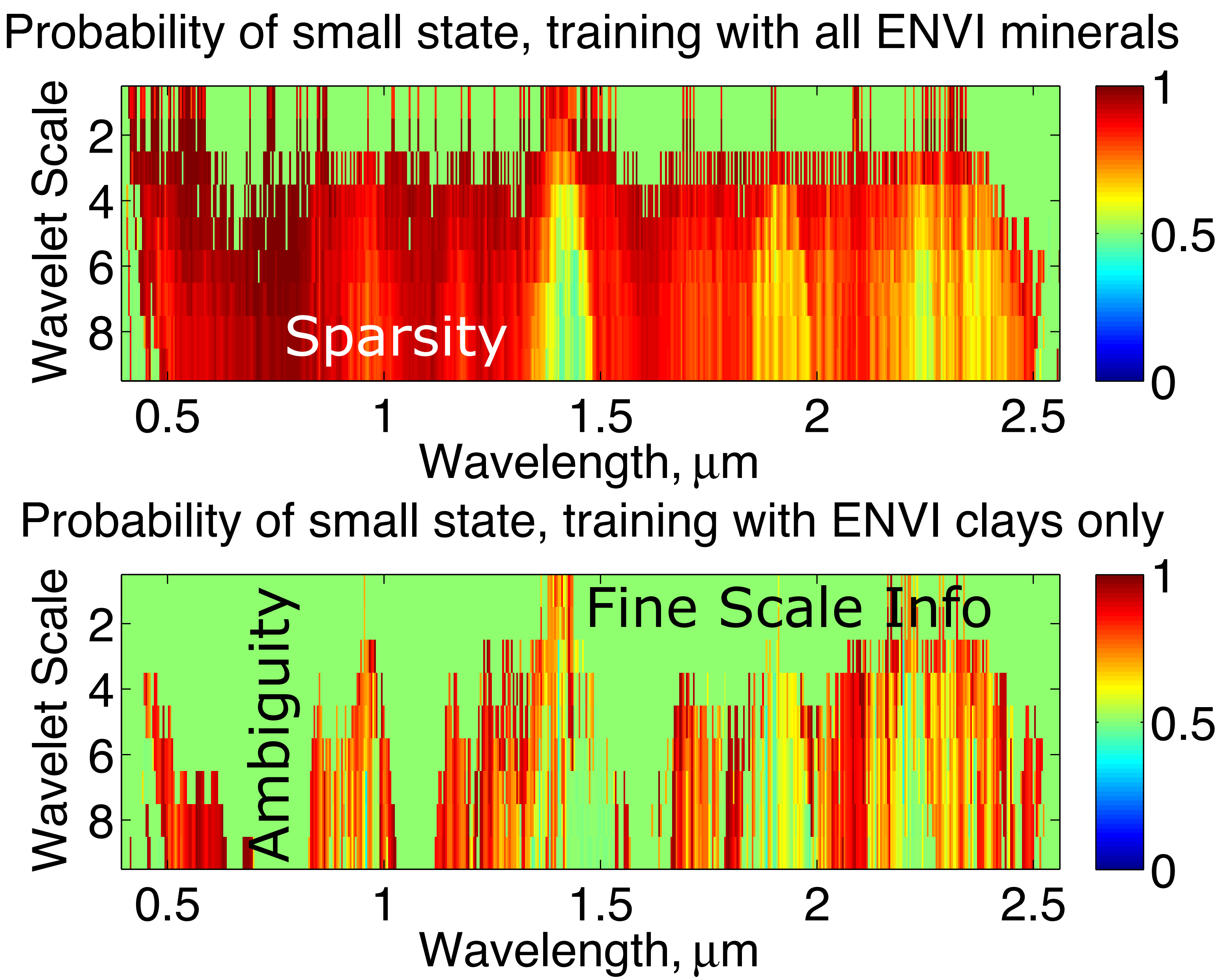
- Statistical modeling of coefficients across spectral sample provides ***measures of relevance*** of bands/smooth regions
- Model parameters can provide ***"map"*** of relevant scales, spectral bands, etc. for training dataset



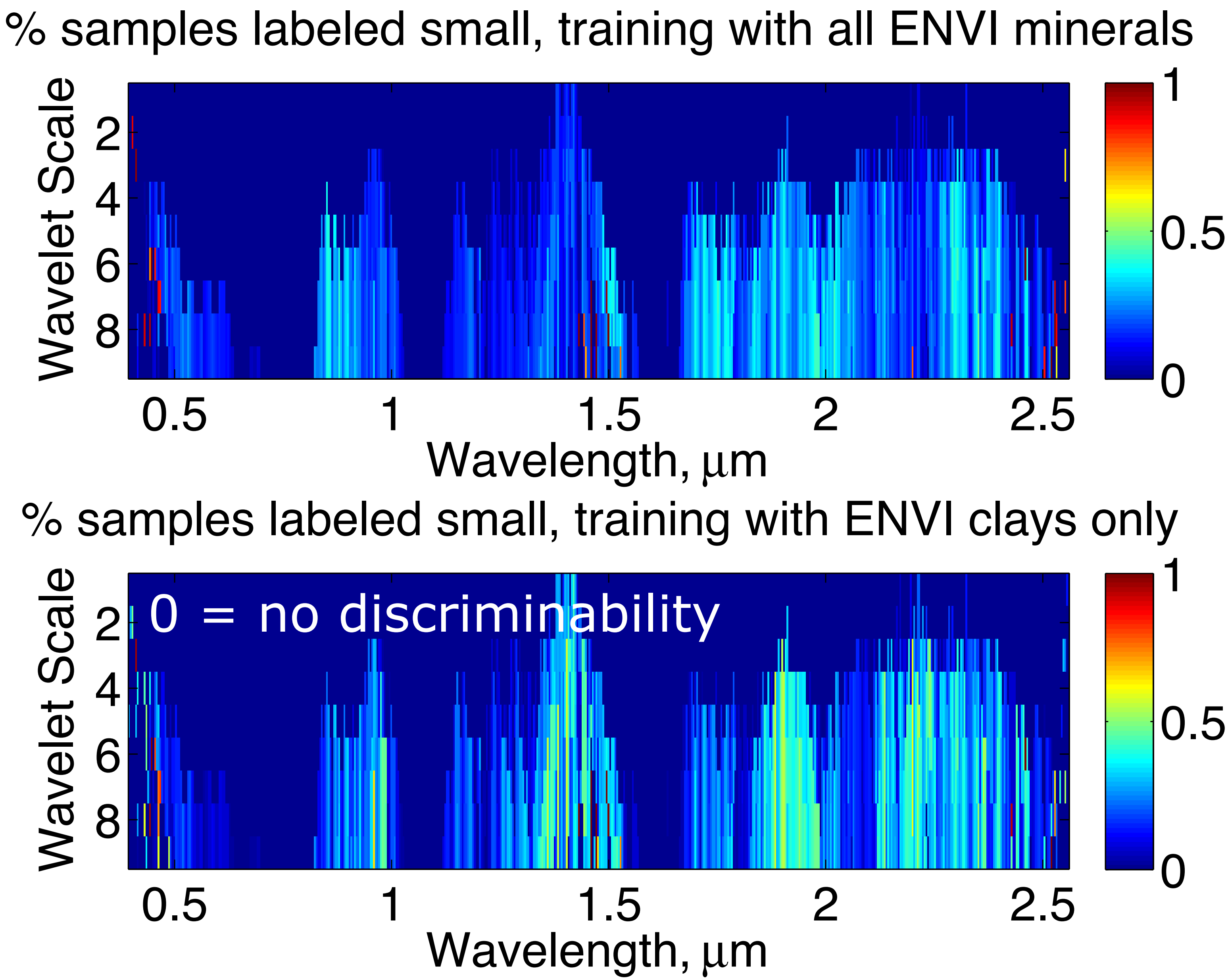
The Power of "Big Data"



The Power of "Big Data"

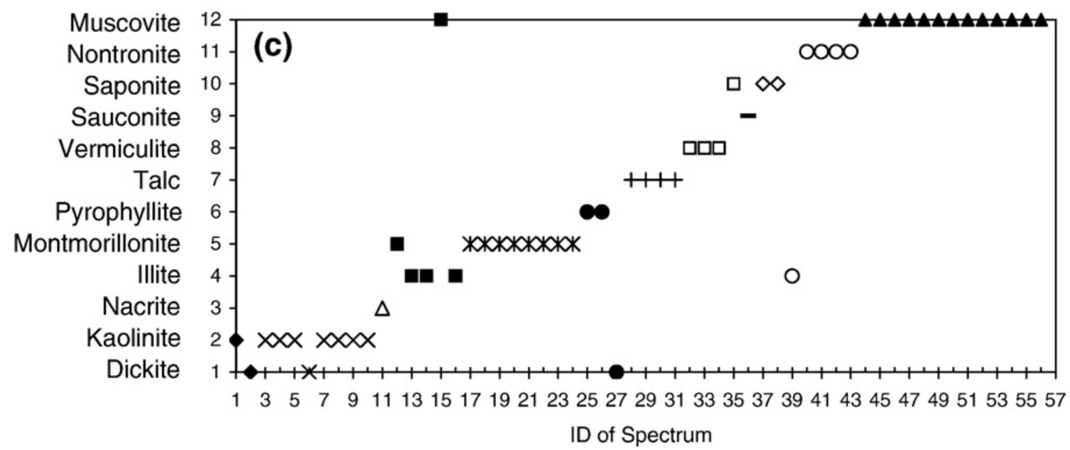


The Power of “Big Data”

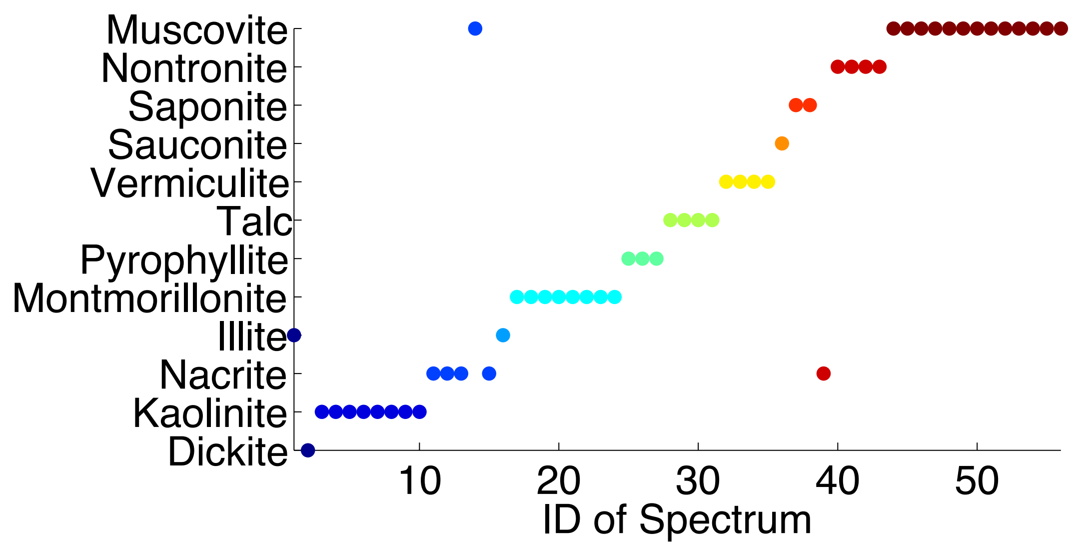


Example: Mineral Classification

- Same example as before, but ***subset of labels*** selected according to three "discriminability" criteria
- For all metrics used, classification performance ***matches*** that obtained with all labels (95% success rate)



[Rivard et al., 2008] 89%



NHMC 95%

Conclusions

- Goal: design hyperspectral signal models and features that can capture ***semantic information*** used by practitioners in remote sensing
  - ***relevance*** of absorption bands in tasks, e.g., classification
  - multiscale analysis studies a variety of spectral features
  - ***robustness*** to fluctuations in shape and location of bands
- Stochastic models (Non-Homogeneous Markov Chain) enable ***robust*** identification of ***relevant features***
  - adaptive sampling, spectral sampling rate adjustments
  - identify non-informative absorption bands, universal features
- ***Future work***:
  - Hyperspectral image applications: segmentation, unmixing, ...
  - Study robustness to signature fluctuations (lab & field datasets)

http://www.ecs.umass.edu/~mduarte
mduarte@ecs.umass.edu