SAND2017-6433C

Advanced Data Analytics in Analytical Electron Microscopy

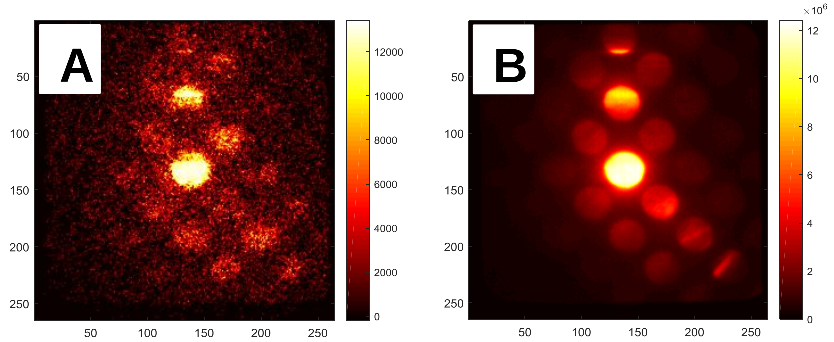
Paul G. Kotula[1], Mark H. Van Benthem[1], Henning Ryll[2], Martin Simson[3], and Heike Soltau[3]

[1]Sandia National Laboratories, Albuquerque, NM, USA

[2]PNSensor GmbH, Otto-Hahn-Ring 6, 81739 Munich, Germany

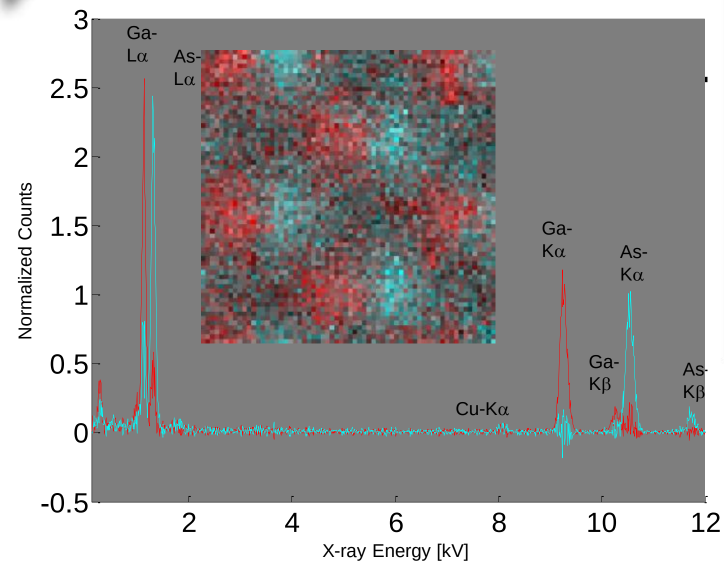
[3]PNDetector GmbH, Otto-Hahn-Ring 6, 81739 Munich, Germany

paul.kotula@sandia.gov



Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Materials Characterization Tetrahedron



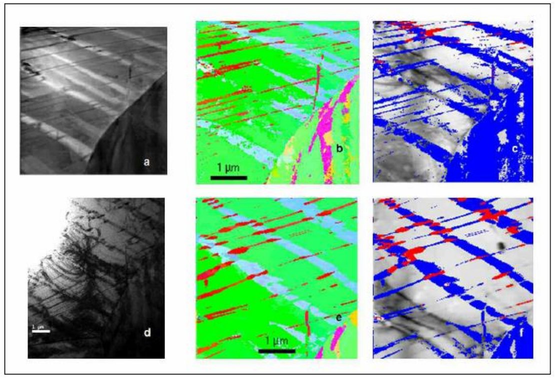
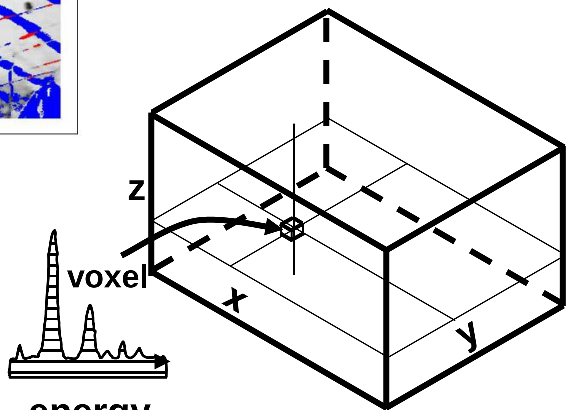
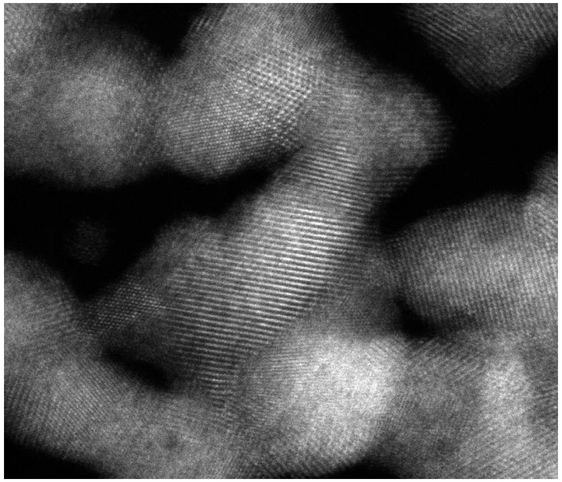


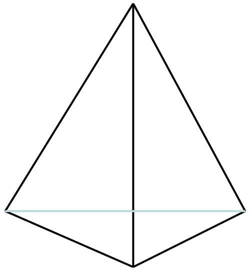
Figure 7:
Austenitic stainless TRIP steel containing three distinct phases: γ fcc matrix, hcp ϵ bands and bcc α' martensite at bands intersections.
(a) Index map coupled to (b) orientation (b) and phase (c) maps acquired without precession.
(d) Brightfield image coupled to orientation (e) and phase (f) maps acquired with precession. Step size 22 nm, precession angle 0.4°, acquisition time 1 hour 9 minutes.





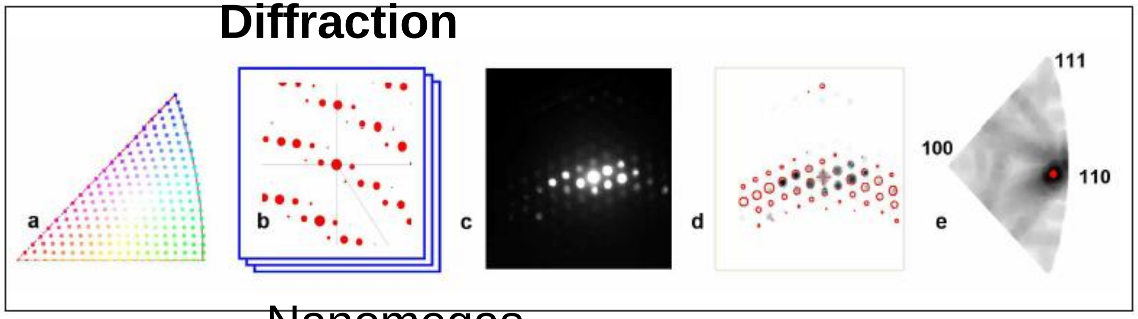
Data Analysis

Imaging



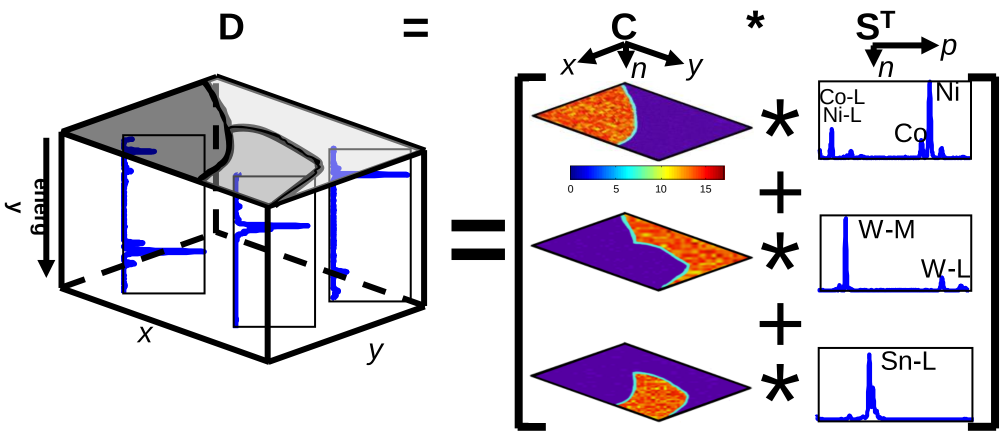
Microanalysis

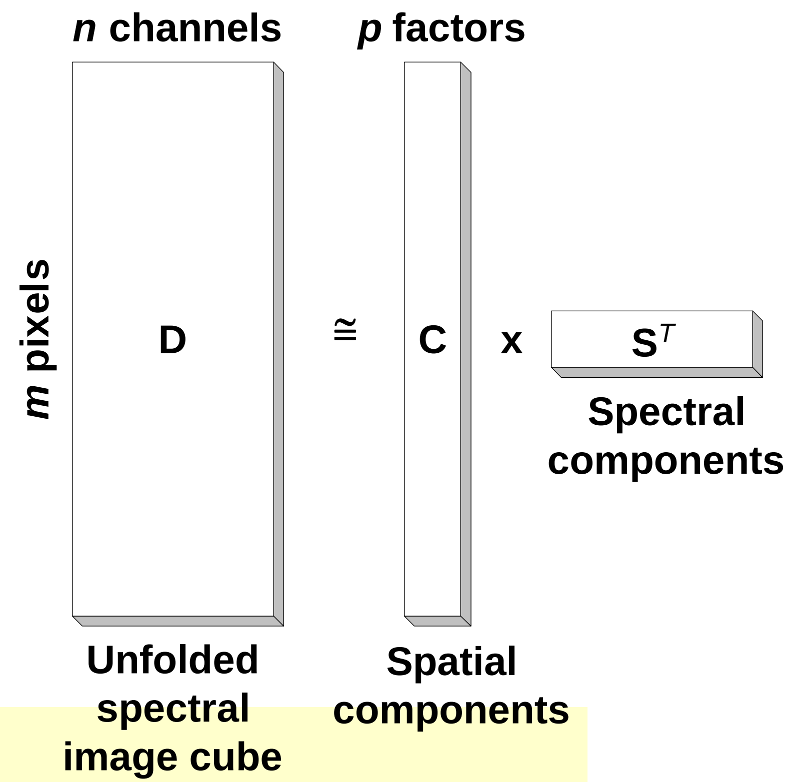
Diffraction



Nanomegas

Multivariate Statistical Analysis (spectroscopy)



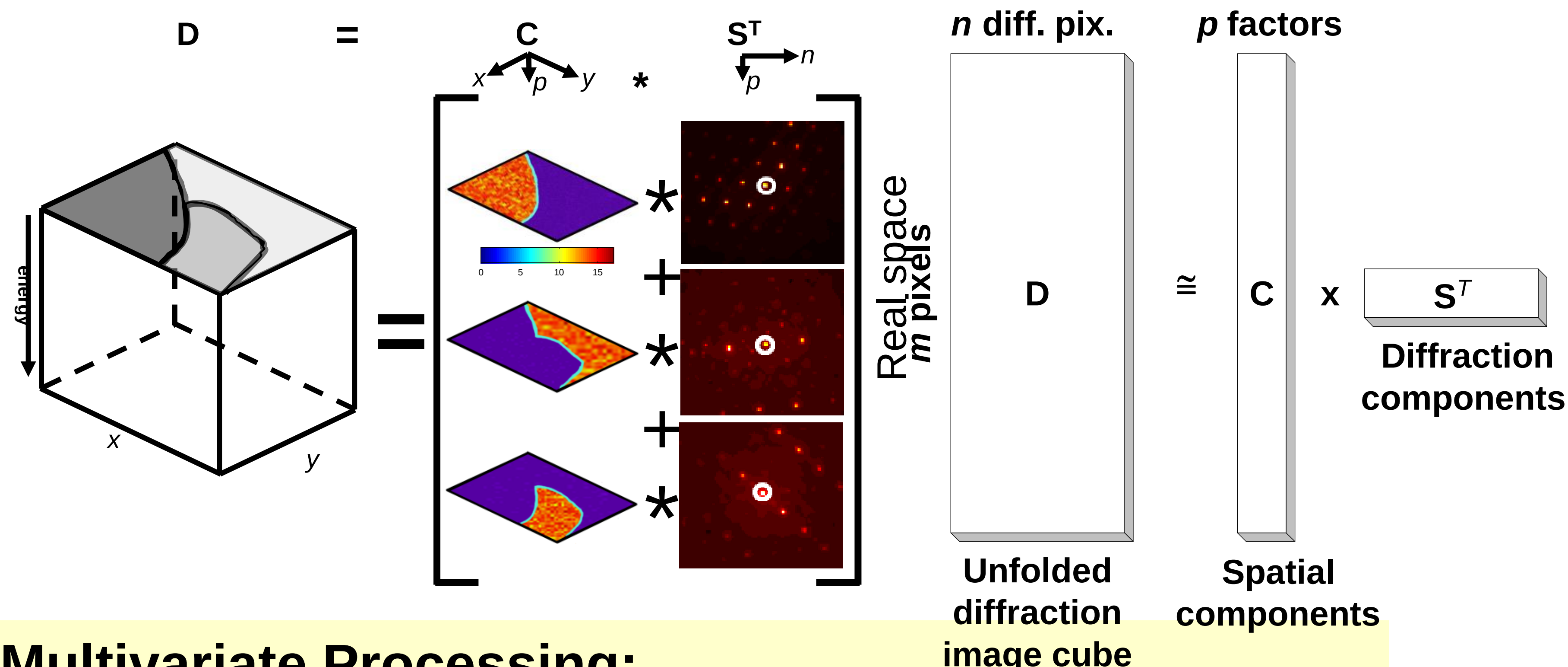


Kotula, P.G., Keenan, M.R. & Michael, J.R. (2003). *Microsc Microanal* **9**, 1–17.

Multivariate Processing:

- **Scale data for Poisson counting statistics**
- **Determine the number, p, of components to keep**
- **Factor the data matrix (D) into C and S**
- **Inverse scale the components**

Multivariate Statistical Analysis (diffraction)



Multivariate Processing:

- **Scale data for noise…do we know this?**
- **Determine the number, p, of components to keep**
- **Factor the data matrix (D) into C and S**
- **Inverse scale the components**

Sandia National Laboratories

What are the basic steps of MSA for spectroscopic data?

- Scale data for non-uniform noise*
  - Assumption here-we know the noise structure in these counting experiments
  - Down-weights large variations in intense spectral or image features which are due to noise
  - Rank 1 approximation to the noise
    - In the image domain divide by the square-root of the mean image (**G**)
    - In the spectral domain divide by the square-root of the mean spectrum (**H**)
    - Essentially the same answer as maximum likelihood methods with but far less computational complexity**

$$\tilde{\mathbf{D}} = \mathbf{GDH}$$

- Factor the data: Keenan, M.R., Multivariate analysis of spectral images composed of count data, in Techniques and applications of hyperspectral image analysis, H. Grahn and P. Geladi, Editors. 2007, John Wiley & Sons: Chinchester.

- Inverse-scale for noise

- For diffraction data we don't necessarily know the noise *a priori*

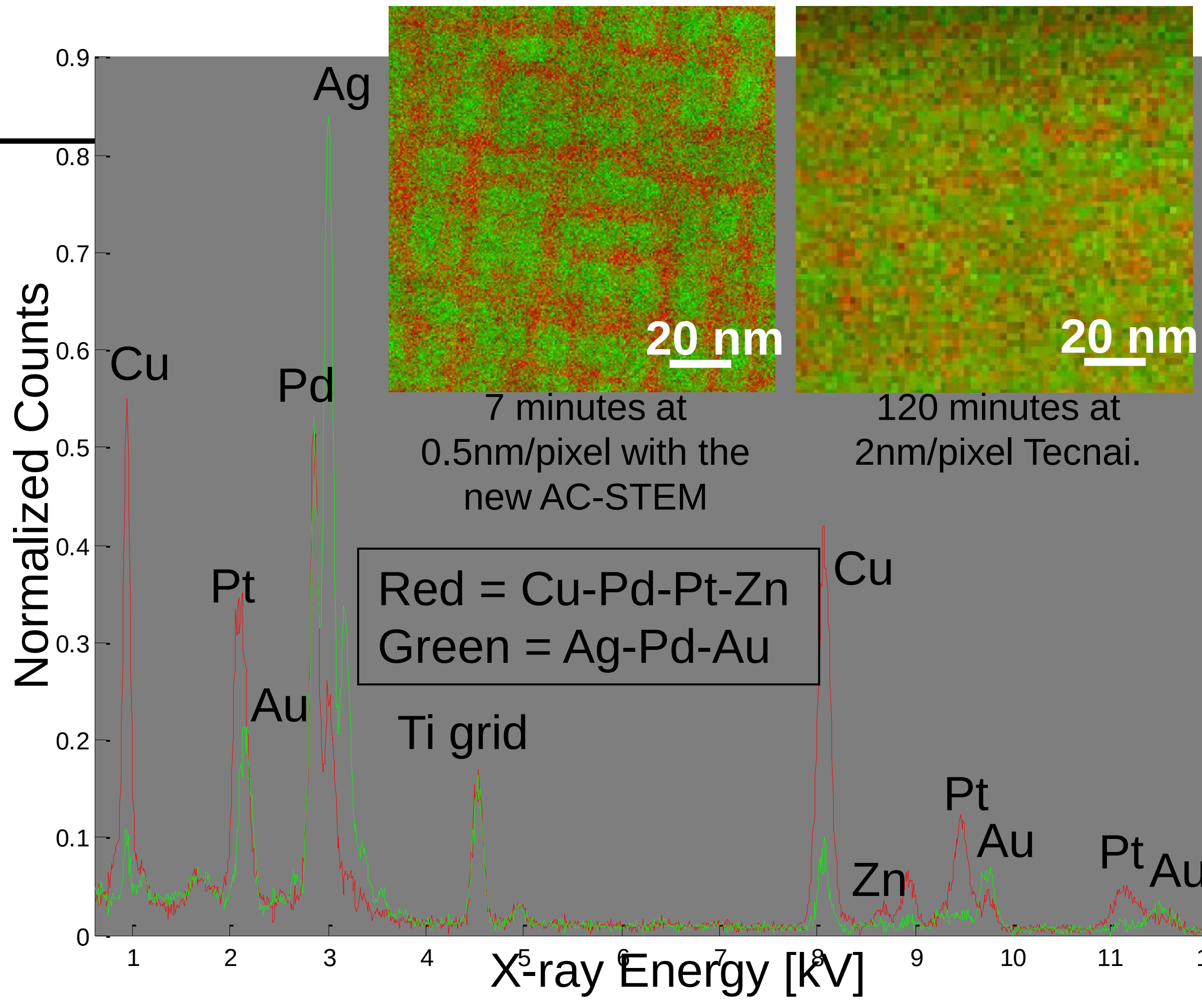
*M.R. Keenan and P.G. Kotula, *Surf. Int. Anal.* **36** (2004) 203-212
M.R. Keenan, J. Vac. Sci. Tech. A **23 [4] (2005) 746-750

Sandia National Laboratories

Sub-nm microanalysis of electrical contact materials

Titan G2 80-200 Cs probe corrected with ChemiSTEM (SuperX)
MSA analysis

Paliney 7, electrical contact material nanometer-scale spinodal decomposition.



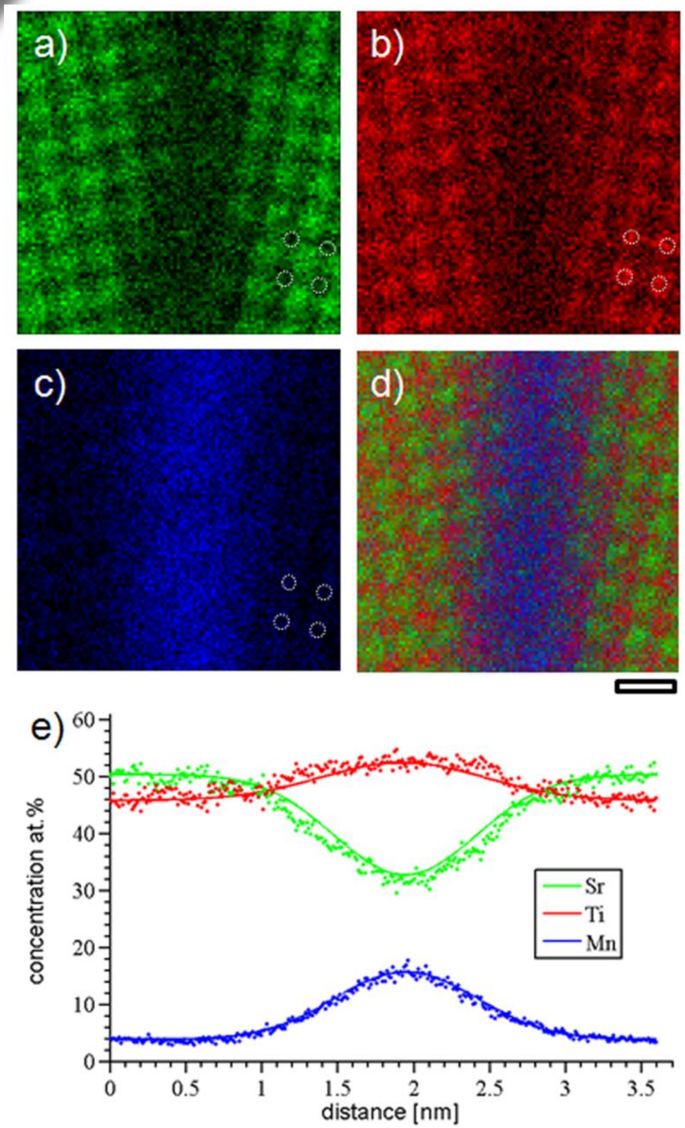
The new AEM is at least **70x** better than our older AEM (Philips/FEI Tecnai F30-ST)

D.F. Susan, Z. Ghanbari, P.G. Kotula, J.R. Michael & M.A. Rodriguez,
Metall. and Mat. Trans. A 45A (2014). DOI: 10.1007/s11661-014-2334-x

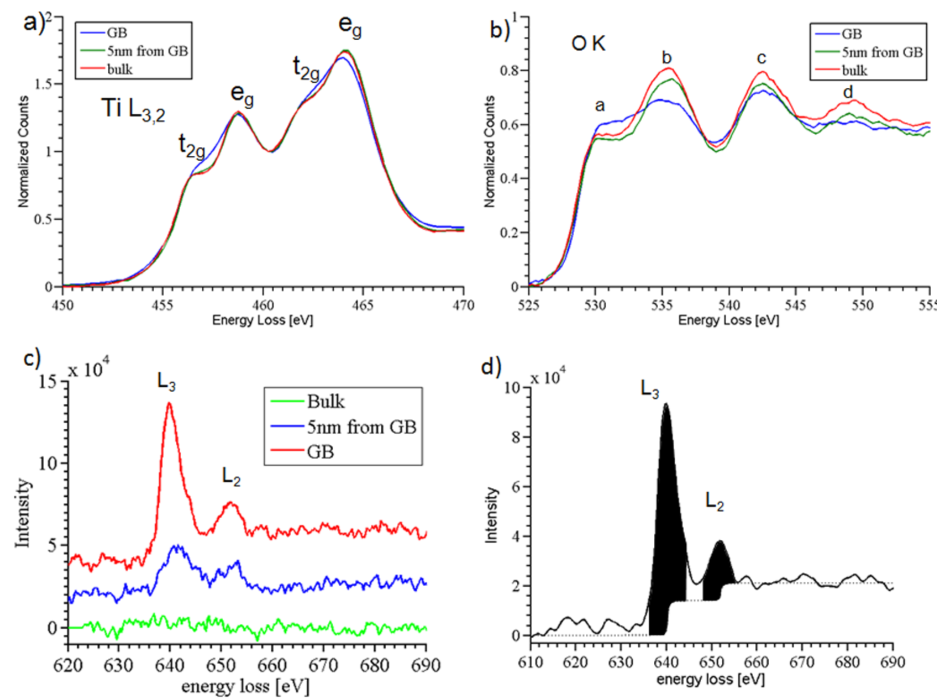
Sandia National Laboratories

Analysis of Mn-doped STO Σ=13 Boundary

∑13 (510)/[001]



Quantitative EDS and EELS



FWTM enrichment at the boundary less than 2nm

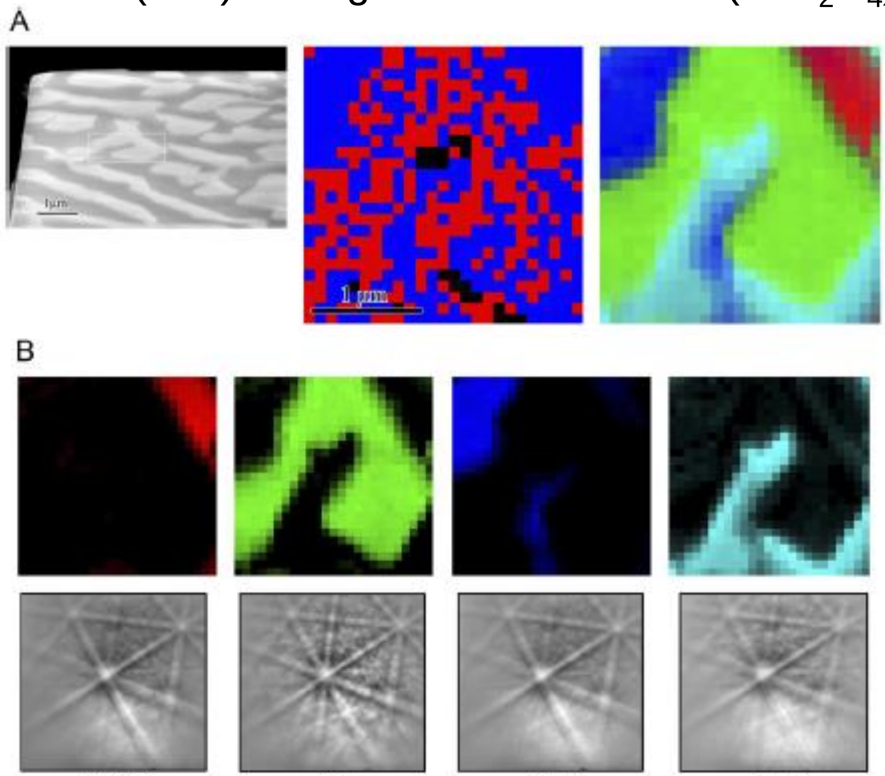
Mn^{+2} at boundary
Mn^{+4} in bulk near boundary (substitutional with Ti)

H. Yang, P.G. Kotula, Y. Sato, Y. Ikuhara, N.D. Browning. *Materials Research Letters* (2013). DOI: 10.1080/21663831.2013.856815

Sandia National Laboratories

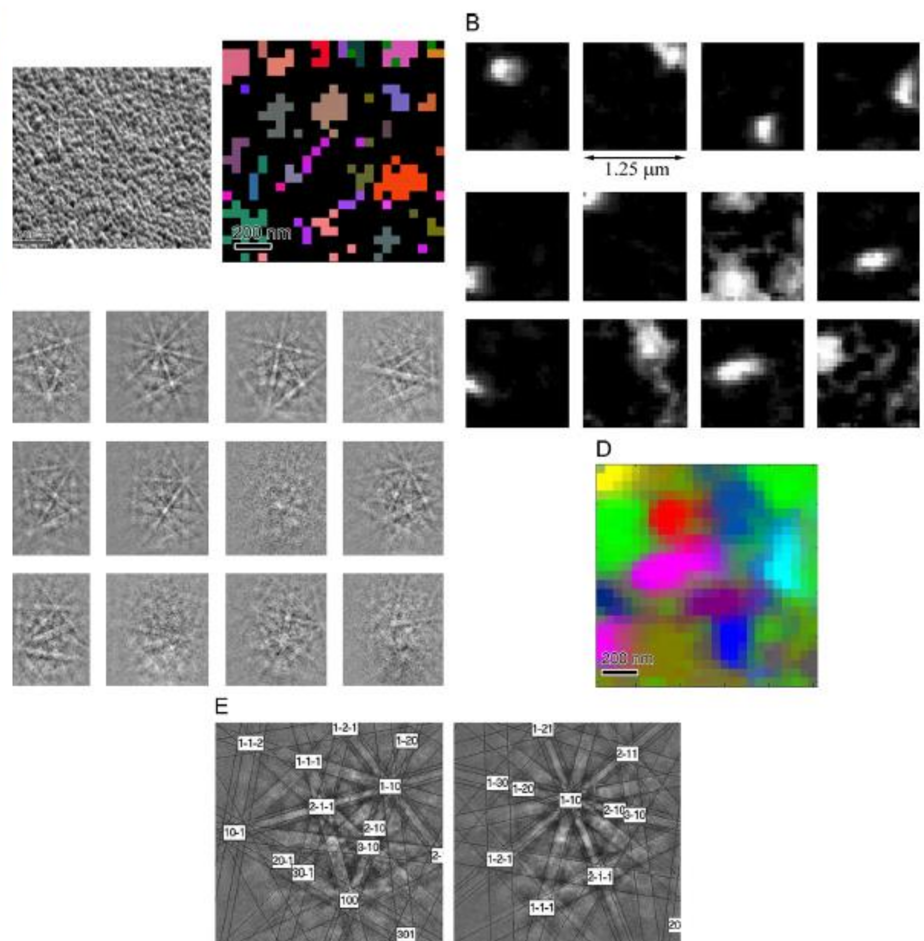
EBSD/MSA*

Fm3m (NiO) distinguished from Fd3m ($NiAl_2O_4$)



*Hough transforms used for MSA versus the raw patterns

Fine Al grains



L.N. Brewer, P.G. Kotula, J.R. Michael, (2008) *Ultramicroscopy* **108**, 567-578.

Sandia National Laboratories

Direct Electron camera: Re-thinking a STEM detector

Traditional STEM detectors are annular (LADF/HAADF) or round (BF)

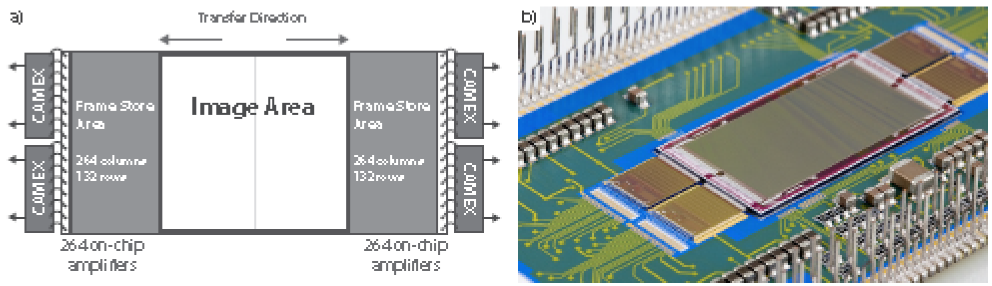
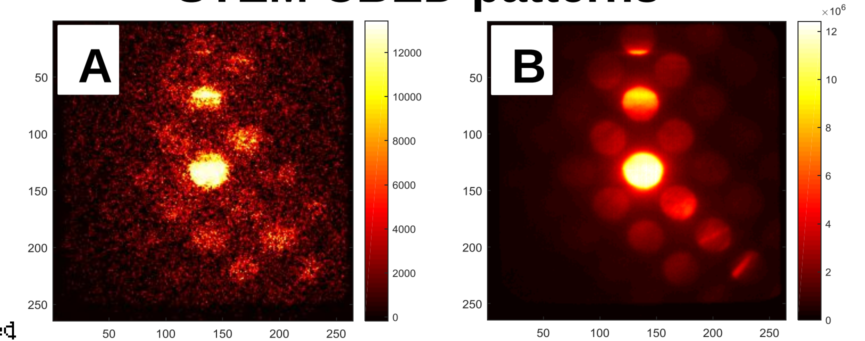


Figure 2. (a) Schematic view of the pnCCD demonstrating its multi parallel readout scheme. The collected charge in the image area is quickly transferred in two directions into the frame store area. Each column is read out via its own on-chip amplifier and connected to individual inputs of one of four readout ASICs called CAMEX. (b) Image of the pnCCD and the four CAMEX mounted on a ceramic holder.

STEM CBED patterns

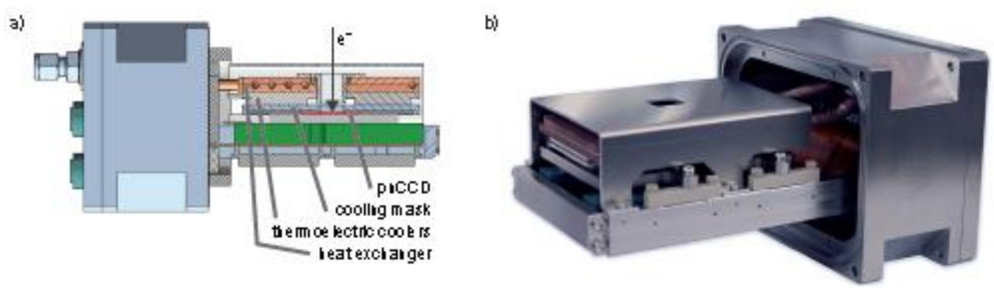


1 msec (1 frame)

20 msec (replicates summed)

Table 1. Specifications overview of the pnCCD used for the (S)TEM-camera.

physical pixel size	$48 \times 48\ \mu m^2$
pixels in image area	264×264
image area	$12.7 \times 12.7\ mm^2$
readout rate	
full frame	1,150 fps
$4\times$ binning	4,000 fps
SNR (see section 4.2)	up to 380:1

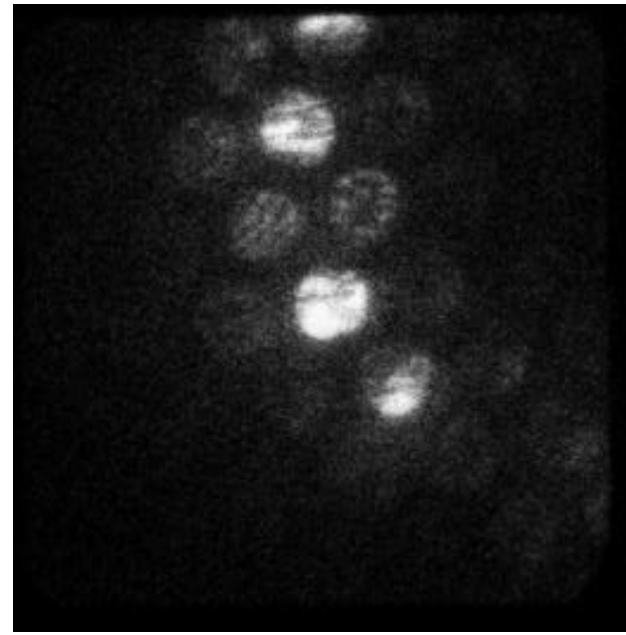


H. Ryll, et al., (2016) "A pnCCD-based, fast direct single electron imaging camera for TEM and STEM." J. Inst.
http://dx.doi.org/10.1088/1748-0221/11/04/P04006

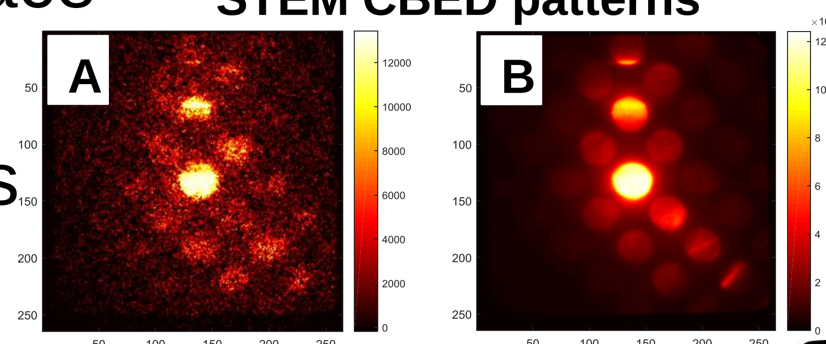
Sandia National Laboratories

STEM-CBED and MSA…A Tale of Big Data

- Titan G2 80-200 STEM
- pnCCD mounted below the projection chamber
- 4.8 mrad convergence angle in probe with 30 pA
- 256x256 real-space pixels
- 264x264 reciprocal-space pixels
- 4×10^9 data elements as summed
- No “small” dimension



STEM CBED patterns



1 msec (1 frame)

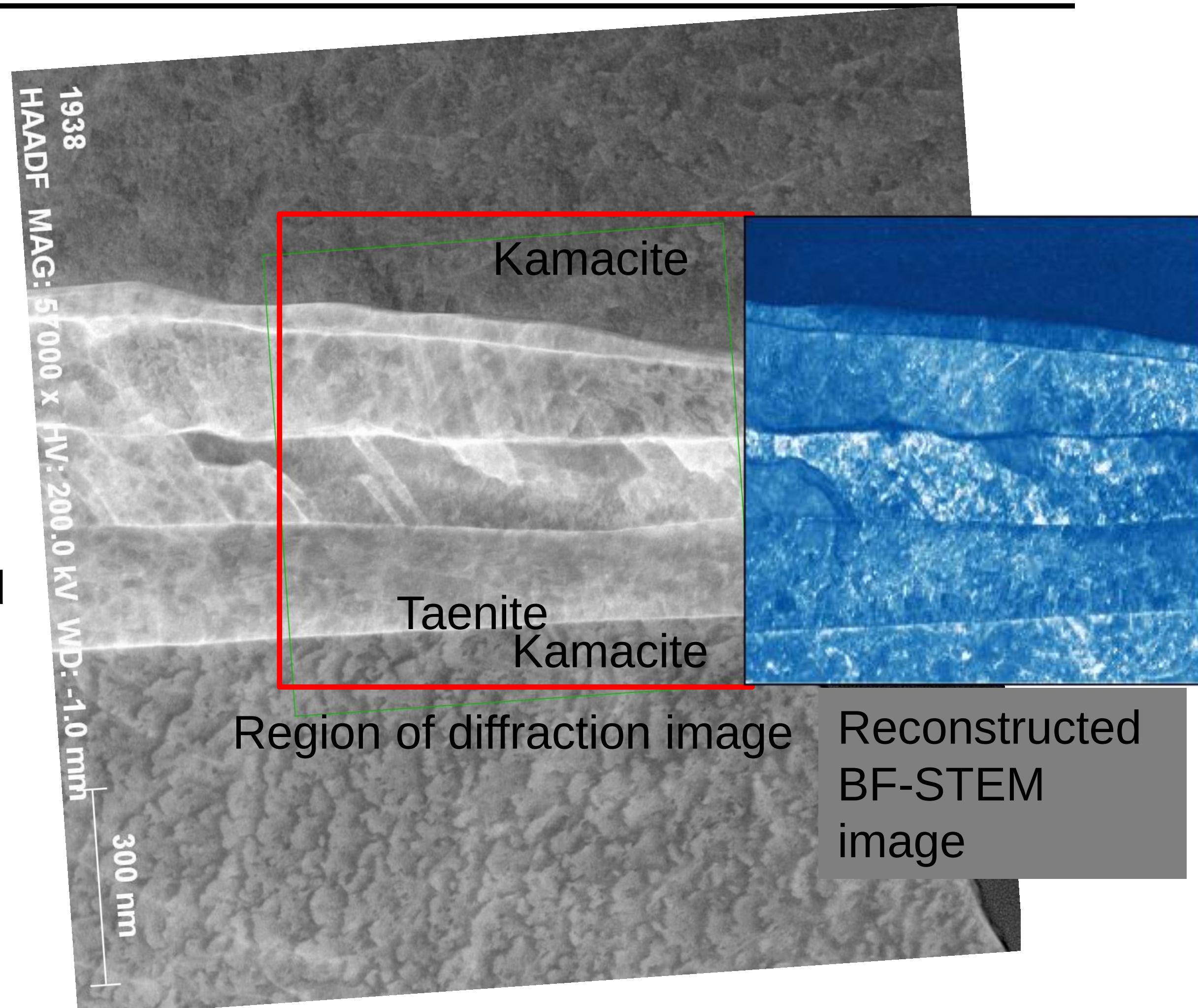
20 msec (replicates summed)

Sandia National Laboratories

STEM-CBED and MSA…A Tale of Big Data

Arltunga meteorite*, Fe-Ni particle

- Data set is 182.7 Gbytes
  - 256x256 real-space pixels
  - 264x264 reciprocal-space pixels
  - 20 msec/pixel total recorded as 20, 1msec diffraction patterns
  - Stored as 2-byte signed integers on a SSD
  - We used the replicate diffraction patterns to determine the noise variance for scaling the data prior to MSA

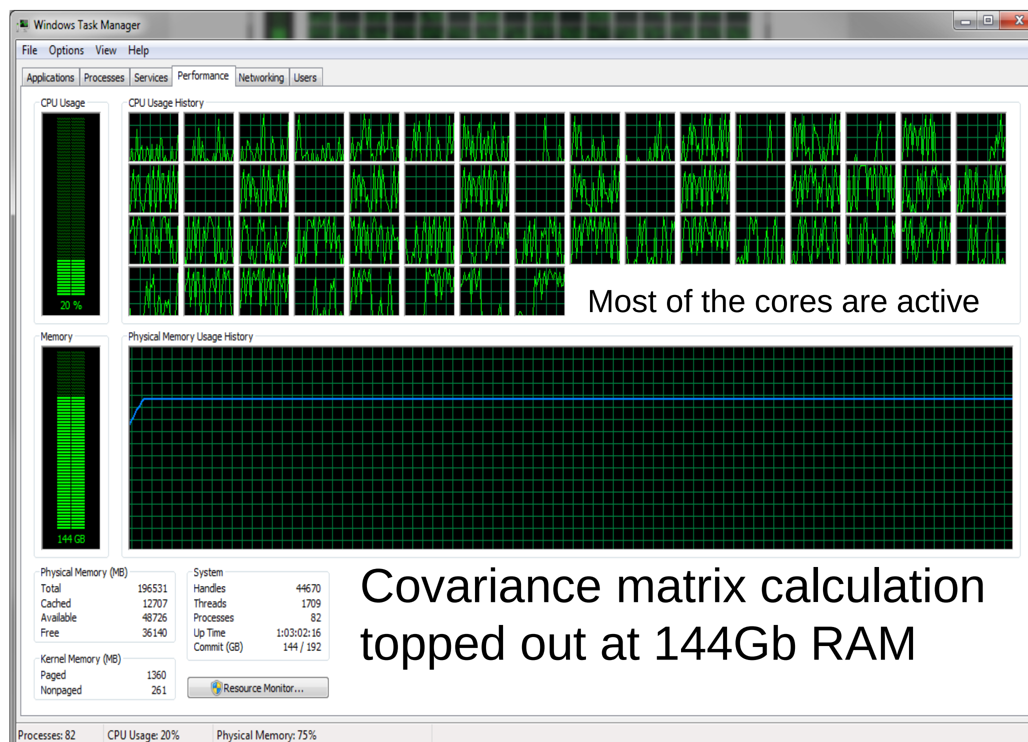


HAADF

*Sample courtesy Joe Goldstein (dec)

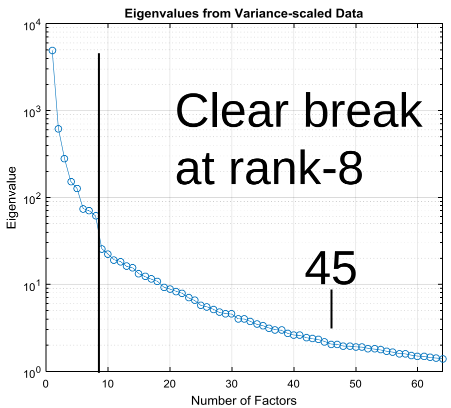
STEM-CBED MSA about 2 hours (non-optimized)

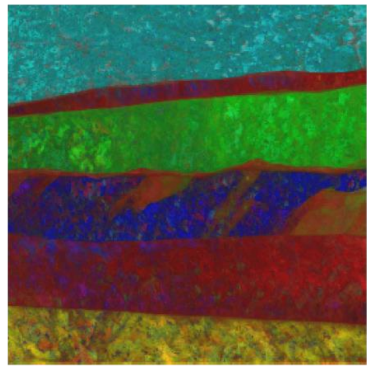
- The covariance matrix is 65536x65536!
  - We typically solve very large problems but one dimension (spectral) is small and therefore calculating the covariance matrix is efficient
    - 16M pixels by 4096 channels has more data elements
- Brute force, this can be solved with a **~$14k computer (~$5k today)**
  - Intel XEON 2, two 28-core processors with 192 Gbytes RAM, Matlab 2016A (64-bit)
  - ~2 hours total calculation time
  - Reading data and calculating the variance images, **93 min.**
  - Calculating the covariance matrix, **13 minutes**
  - Eigenanalysis, **21 minutes**. PCA-Varimax (rank=8) **45 sec.**
  - Data read time could be optimized greatly (e.g., HDF-5)



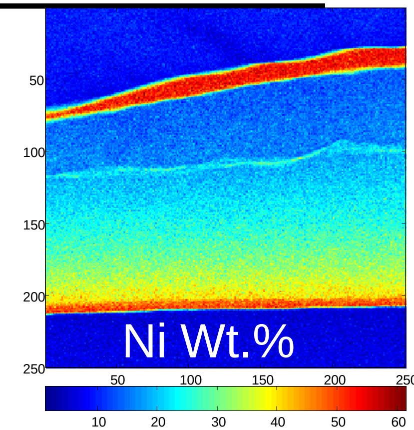
STEM-CBED and MSA…Real-space simplicity

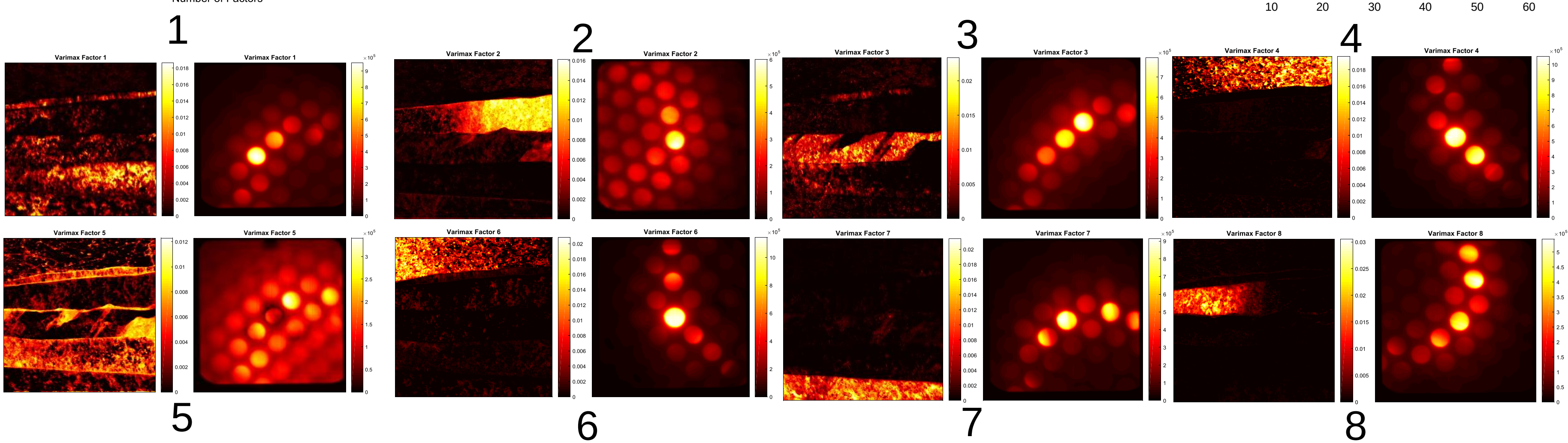
38 Gbyte in, 9 Mbyte after MSA, Compression factor of ~21000





Red = 1+5 (fcc)
Green = 2+8 (fcc)
Blue = 3 (fcc)
Cyan = 4+6 (bcc)
Yellow = 7 (bcc)

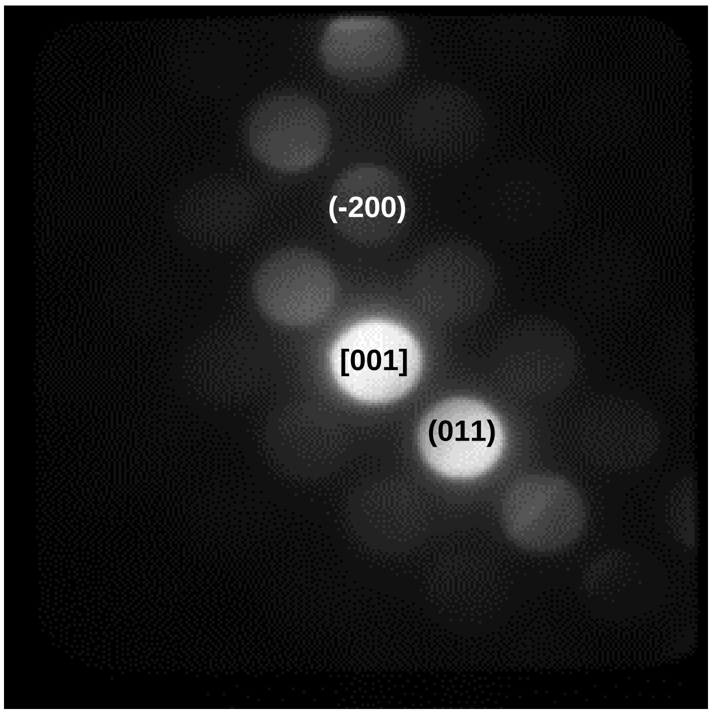


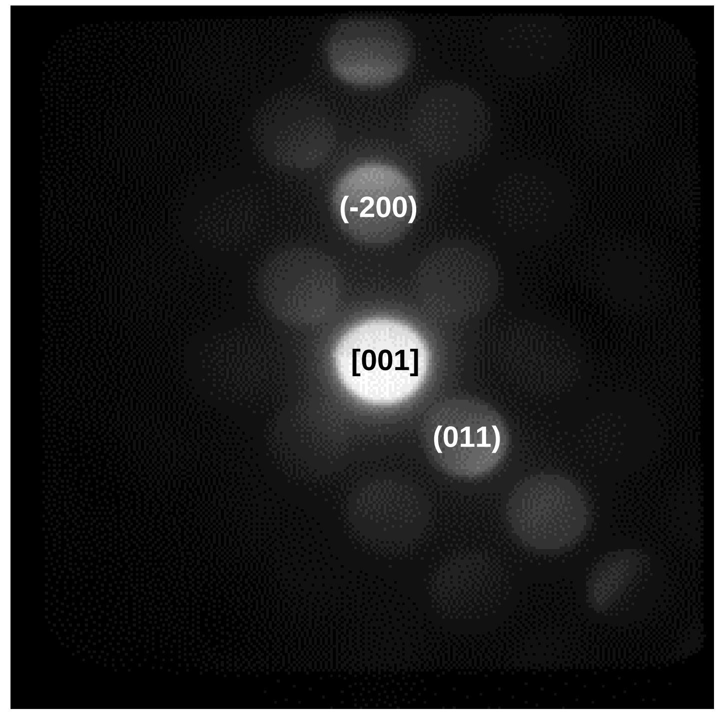


First-pass MSA breaks out largely by crystallographic phase/orientation

9-component model

Top BCC phase described by at just two factors



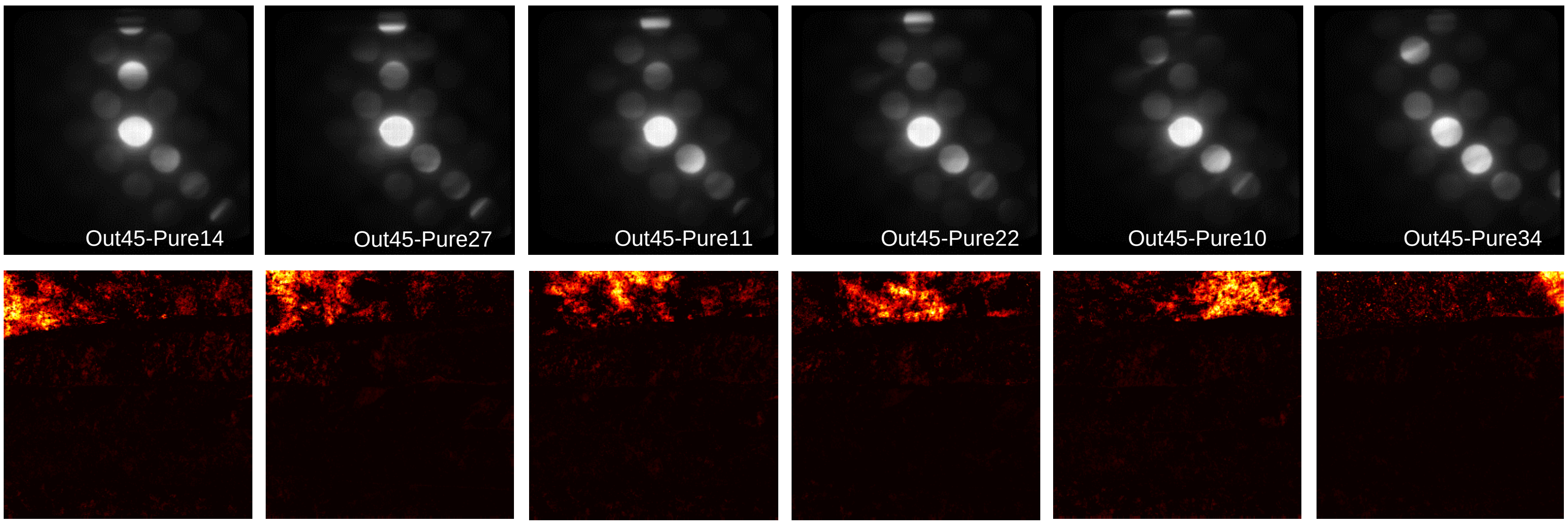


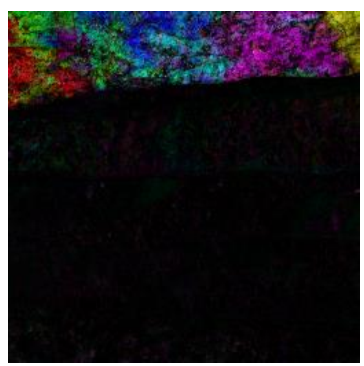
Both BCC near [001] slightly different orientations

Sandia National Laboratories

45-component model

Top BCC phase described by at lease six factors

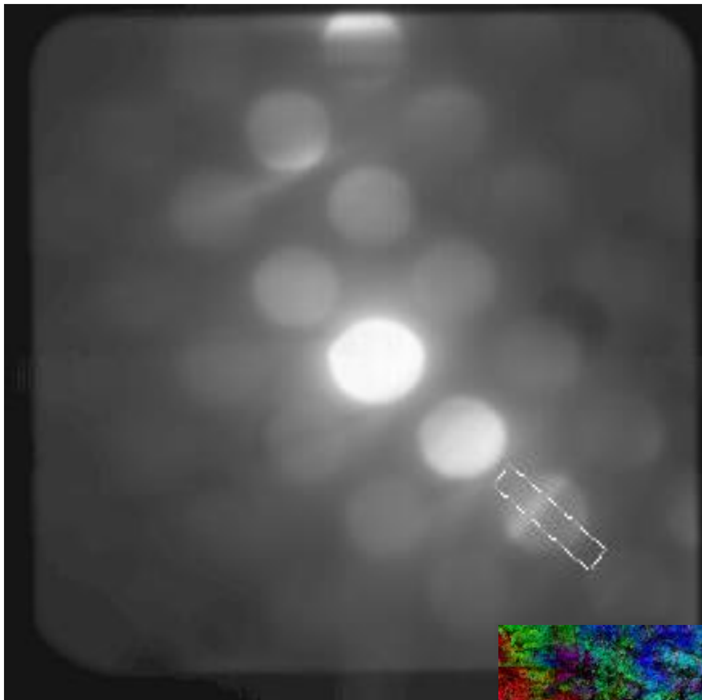


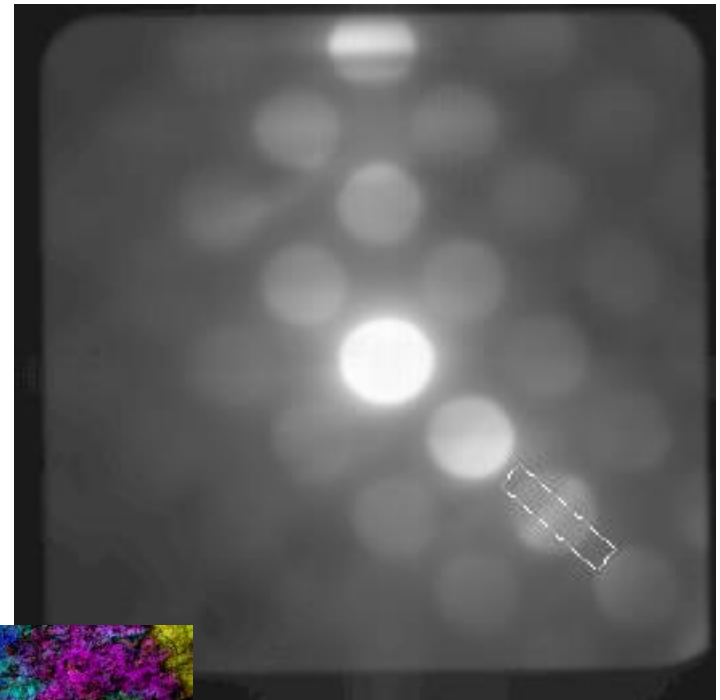


Red = 14
Green = 27
Blue = 11
Cyan = 22
Magenta = 10
Yellow = 34

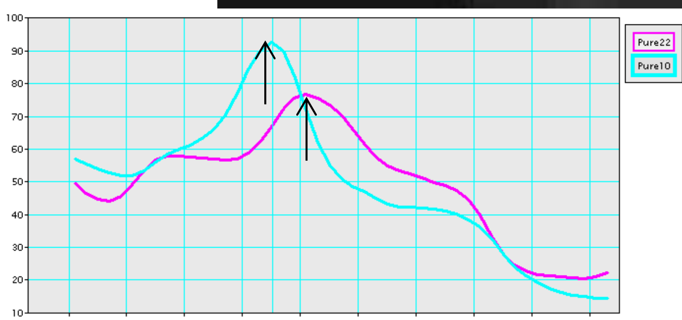
Sandia National Laboratories

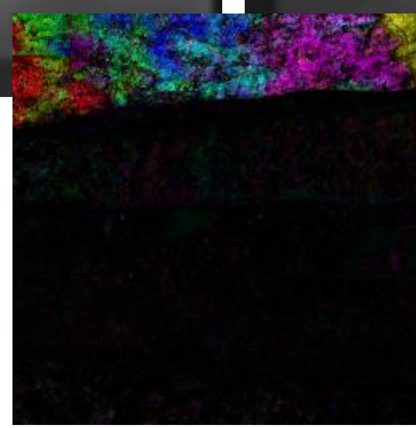
10





22





Red = 14
Green = 27
Blue = 11
Cyan = 22
Magenta = 10
Yellow = 34

Kikuchi Band shift of 4 pixels corresponds to 0.03° tilt of the crystal

Sandia National Laboratories

Conclusions

- Data analytics makes it possible to make sense of large, noisy redundant data sets by reducing their dimensionality
- Relatively straight forward for spectral images
- More complicated for image or diffraction series
- Direct electron cameras can replace conventional STEM detectors and allow for retrospective analysis.
  - Replicate diffraction patterns allow us to determine the noise variance from the data itself
  - Rank estimation determines the level of detail (8 versus 45 shown)