N/A

How to Avoid Unexpected Artifacts from Multivariate Statistical Analysis on STEM Spectrum-Imaging Datasets

Ishizuka K.[1], Watanabe M.[2]

[1]HREM Research Inc., Higashimatsuyama, Japan, [2]Dept of Materials Science and Engineering, Lehigh University, Bethlehem, USA

Email of the presenting author: ishizuka@hremresearch.com

The latest aberration-corrected scanning transmission electron microscope (STEM) makes possible to perform routinely not only atomic-scale imaging but also chemical analysis via electron energy-loss spectrometry (EELS) and X-ray energy dispersive spectrometry (XEDS) [e.g. 1]. In combination with the latest hardware, the advances in the recent software developments allow us to acquire large-scale datasets such as multidimensional image series and spectrum images (SIs). Therefore, it is challenging to deal with the large-scale datasets, e.g. extraction of unknown features and estimation of dominant trends. If the datasets were relatively noisy, which is very common for atomic-resolution EELS/XEDS SIs, data analysis would be much harder tasks. Multivariate statistical analysis (MSA) is one of efficient approaches to analyse the large-scale datasets in terms of feature identification and extraction.

Principal component analysis (PCA) is one of the MSA techniques [2]. Since a use of PCA is relatively straightforward, PCA has been applied to SIs as data-mining and noise-reduction tools [e.g. 3]. The PCA tries to explain the data variation (variance) as much as possible using a small number of the components. Here, the signal itself of course contributes the data variation. However, a small amount of signal will be buried with the whole random noise. Therefore, despite that the PCA approach is very efficient and useful, it may create unexpected artifacts especially in higher noise conditions [4] (Figure 1). Since these artifacts might mislead results, it is essential to avoid such artifacts. There may be two approaches to improve the PCA sensitivity: (1) reduction of random noise and (2) enhancement of true variations. The former requires modifications in experimental conditions (higher currents and longer acquisitions). Conversely, the latter can be achieved by PCA analysis to divided small segments within a SI, which is called the local PCA approach (Figure 2). The division can be made spatially and spectrally. The spatially local PCA will be especially useful to detect segregated element in the matrix. The spectrally local PCA is useful to detect a weak signal, if the weak signal is spectrally separated from the strong signal. Especially the spectrally local PCA is useful for EELS Sis, since the background intensity varies significantly. In this study, advantages of the local PCA approach will be addressed.

[1] S.J. Pennycook & P.D. Nellist ed. Scanning Transmission Electron Microscopy: Imaging and Analysis, Springer, NY, (2011).

[2] E.R. Malinowski, Factor Analysis in Chemistry, 3rd ed., Wiley, New York, (2002).

[3] M Watanabe et al., Microscopy and Analysis, 23, Issue 7, (2009), 5-7.

[4] S. Lichtert & J. Verbeeck, Ultramicrosc., 125 (2013), 35-42

Acknowledgement: The authors acknowledge J. Verbeeck for providing the simulated BN test data. M.W. wishes to acknowledge financial support from the NSF through grants DMR-0804528 and DMR-1040229.

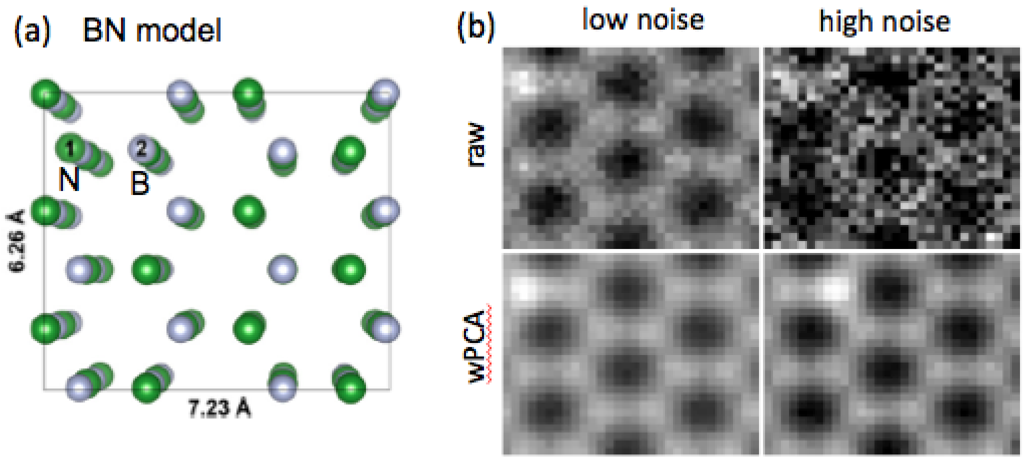

Fig. 1: Failure of PCA [4]. (a) BN model, where there are one excess N or B atom at the positions 1 and 2, respectively. (b) Untreated and weight PCA N element maps for low and high noise. Note that the N map of high noise shows higher intensity at 2 than 1.

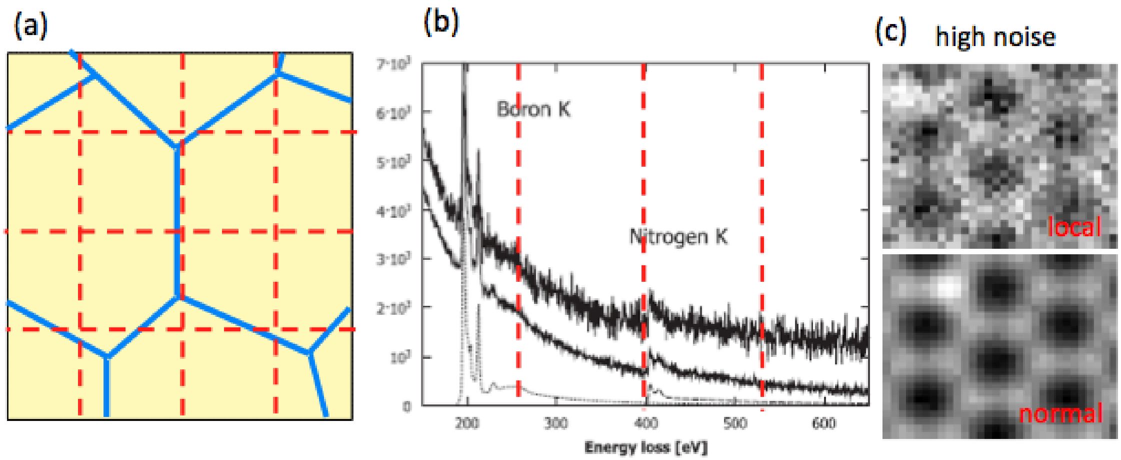

Fig. 2: (a) and (b): Spatially local PCA and spectral local PCA, respectively. (c): N maps reconstructed by the spectrally local PCA shows higher intensity at the position 1 even for the high noise case contrary to the normal PCA.