Improving Quantitative EDS Chemical Analysis of Alloy Nanoparticles by PCA Denoising: Part I, Reducing Reconstruction Bias

Murilo Moreira, Matthias Hillenkamp, Giorgio Divitini, Luiz H G Tizei, Caterina Ducati, Monica Cotta, Varlei Rodrigues, Daniel Ugarte

MICROSCOPY AND MICROANALYSIS

Microscopy Society of America

CAMBRIDGE UNIVERSITY PRESS

Improving Quantitative EDS Chemical Analysis of Alloy Nanoparticles by PCA Denoising: Part I, reducing reconstruction bias.

Journal:	*Microscopy and Microanalysis*
Manuscript ID	MAM-21-338.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	21-Oct-2021
Complete List of Authors:	Moreira, Murilo; UNICAMP, DFA-IFGW Hillenkamp, Matthias ; Université Claude Bernard Lyon 1, Institut lumière matière Divittini, Giorgio; University of Cambridge, Materials Science and Metallurgy Tizei, Luiz; Laboratoire de Physique des Solides Ducati, Caterina; University of Cambridge, Materials Science and Metallurgy Cotta, Mônica; UNICAMP, DFA-IFGW Rodrigues, Varlei; UNICAMP, DFA-IFGW Ugarte, Daniel; UNICAMP, DFA-IFGW; UNIV Estadual de CAmpinas
Keywords:	Nanoparticles, Transmission Electron Microscopy-TEM, Energy dispersive X-ray Spectroscopy-EDS, Machine Learning, quantitative chemical analysis
Abstract:	Scanning Transmission Electron Microscopy (STEM) is a crucial tool for nanoscience, achieving sub-nanometric spatial resolution in both image and spectroscopic studies. This generates large datasets that cannot be analyzed without computational assistance. The so-called "machine learning" procedures can exploit redundancies and find hidden correlations. Principal Component Analysis (PCA) is the most popular approach to denoise data by reducing data dimensionality and extracting meaningful information; however, there are many open questions on the

	accuracy of reconstructions. We have used experiments and simulations to analyze the effect of PCA on quantitative chemical analysis of binary alloy (AuAg) nanoparticles using energy dispersive x-ray spectroscopy (EDS). Our results demonstrate that it is possible obtain very good fidelity of chemical composition distribution when the signal-to-noise ratio exceeds a certain minimal level. Accurate denoising derives from a complex interplay between redundancy (data matrix size), counting noise and noiseless data intensity variance (associated to sample chemical composition dispersion). We have suggested several quantitative bias estimators and noise evaluation procedures to help in the analysis and design of experiments. This work demonstrates the high potential of PCA denoising, but it also highlights the limitations and pitfalls that need to be avoided to minimize artifacts and perform reliable quantification.

Improving Quantitative EDS Chemical Analysis of alloy Nanoparticles by PCA Denoising: Part I, reducing reconstruction bias.

Murilo Moreira[1], Matthias Hillenkamp[1,2], Giorgio Divitini[3], Luiz H. G. Tizei[4], Caterina Ducati[3], Monica A. Cotta[1], Varlei Rodrigues[1,*], Daniel Ugarte[1,*]

1. *Instituto de Fisica "Gleb Wataghin", Universidade Estadual de Campinas-UNICAMP, 13083-859, Campinas - SP, Brazil*
2. *Institute of Light and Matter, Université Claude Bernard Lyon 1, CNRS, UMR5306, F-69622 Villeurbanne, France*
3. *Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB3 0FS, United Kingdom*
4. *Laboratoire de Physique des Solides, Université Paris-Saclay, CNRS, 91405 Orsay, France*

***Correspondence and requests for materials should be addressed to:**
D.U. (email: dmugarte@ifi.unicamp.br) or to V.R. (email: varlei@ifi.unicamp.br)

Abstract:

Scanning Transmission Electron Microscopy (STEM) is a crucial tool for nanoscience, achieving sub-nanometric spatial resolution in both image and spectroscopic studies. This generates large datasets that cannot be analyzed without computational assistance. The so-called "machine learning" procedures can exploit redundancies and find hidden correlations. Principal Component Analysis (PCA) is the most popular approach to denoise data by reducing data dimensionality and extracting meaningful information; however, there are many open questions on the accuracy of reconstructions. We have used experiments and simulations to analyze the effect of PCA on quantitative chemical analysis of binary alloy (AuAg) nanoparticles using energy dispersive x-ray spectroscopy (EDS). Our results demonstrate that it is possible obtain very good fidelity of chemical composition distribution when the signal-to-noise ratio exceeds a certain minimal level. Accurate denoising derives from a complex interplay between redundancy (data matrix size), counting noise and noiseless data intensity variance (associated to sample chemical composition dispersion). We have suggested several quantitative bias estimators and noise evaluation procedures to help in the analysis and design of experiments. This work demonstrates the high potential of PCA denoising, but it also highlights the limitations and pitfalls that need to be avoided to minimize artifacts and perform reliable quantification.

Key words: Nanoparticles, Machine Learning, denoising, Principal Components Analysis-PCA, quantitative chemical analysis, Transmission Electron Microscopy-TEM, Energy dispersive X-ray Spectroscopy-EDS

1. Introduction

Nanotechnology exploits the unique properties of nanoparticles (NPs) in many different fields such as catalysis, magnetism, plasmonics, etc. (Heiz & Landman, 2007; Odom & Schatz, 2011; Binns, 2014). The electronic properties of bimetallic NPs show a complex behaviour in relation to size and elemental composition; for example, several suggested applications exploit the optimization of either physical properties such as surface plasmons in some optical devices, or chemical reactivity/selectivity in catalysis. Morphological and structural characterization of a NP sample requires the measurement of crystalline structure, shape (rod, sphere, wire, etc.), crystal habit (possible faceting), and size distribution (mean diameter and size dispersion). In the case of multi-elemental nanosystems (i.e. nanoalloys (Alloyeau et al., 2012; Ferrando, 2016)), we must determine the chemical composition (mean one and distribution as a function of size), and possibly the occurrence of chemical inhomogeneity within (the particles (core-shell segregation, Janus distribution, compositional gradients, etc. (Mukherjee et al., 2012, 2015; Lyman et al., 1995)).

Transmission Electron Microscopy (TEM) has made huge progress in the last decades, considering electron optics, detector efficiency, reproducibility, automation, etc. (Williams & Carter, 2009, 2016; Hawkes & Spence, 2019). Scanning Transmission Electron Microscopes (STEMs) are capable of recording the so-called hyperspectral image (HSI) (Pennycook & Nellist, 2011), where an entire analytical spectrum or diffraction pattern can be registered at each image pixel. The HSI approach is currently used to generate huge 3D or 4D data sets (X-ray Energy Dispersive Spectroscopy (EDS), Electron Energy Loss Spectroscopy (EELS), Cathodoluminescence (CL), electron diffraction (ED), scattering distribution, etc.) (Thomas et al., 2015).

A typical HSI datacube may routinely contain 10^7 voxels (100x100 pixels, each containing a spectrum of 1000 channels). The information contained in such massive arrays can be leveraged to a better signal-to-noise ratio (SNR) by applying blind source separation algorithms (Williams & Carter, 2016;

Brown et al, 2020; Cueva et al., 2012). By exploiting redundancies, Machine Learning (ML) methods also are able to efficiently classify data by finding hidden correlations. The most popular algorithm is Principal Component Analysis (PCA) (Jolliffe, 2002, Jolliffe & Cadima, 2016), that provides a number of orthogonal eigenvectors (referred to as "loadings" or "components") encoding the information contained in data variance. The reconstruction data with a few PCA components allows the extraction of meaningful information and reduction of noise (usually called "denoising") (Titchmarsh, 1999; Lichtert & Verbeeck, 2013; Keenan & Kotula, 2004; Kotula & Keenan, 2006; Kotula & Van Benthem, 2015; Potapov & Lubk, 2019). EDS data is particularly well suited for PCA processing, because signal is constituted by well-defined peaks on a low varying background (Lichtert & Verbeeck, 2013). In contrast, EELS display absorption edges and background which may show changing complex shapes as a function of chemistry and sample thickness, so PCA application may be rather complex (Cueva et al., 2012; Lichtert & Verbeeck, 2013).

EDS chemical mapping has become one of the most popular TEM related experiments because the high counting capabilities of Silicon Drift Detector (SDD) has strongly reduced total acquisition time (Watanabe et al., 2010; Schlossmacher et al., 2010). PCA has become a popular easy-to-use tool to denoise and improve qualitatively EDS elemental maps. In contrast, it has scarcely been used for quantitative chemical EDS-STEM analysis (Burke et al., 2006), where it is essential to confirm the PCA reconstruction fidelity (accuracy) and calculate reliable confidence intervals. The user-friendly profile of PCA may erroneously indicate a fault-free procedure; indeed, seriously biased reconstructions (i.e. derived values differ from true noiseless ones) may arise from low SNR datasets. The achievement of successful unbiased denoising is strongly dependent on appropriately designed and executed EDS HSI experiments. Despite the widespread use of PCA processing (Burke et al., 2006; Parish & Brewer, 2010; Rossouw, et al., 2015),

there are still many open questions on noise-related effects and on how to infer if reconstructions are biased.

Alloy NP samples are expected to display a size dispersion and elemental composition variation, making their detailed characterization an extremely challenging task. In this work, we have explored the application of PCA denoising in quantitative analysis of binary alloy NP using EDS HSI; this issue represents a rich, complex ensemble of yet open questions. This investigation requires the scrutiny of the two different aspects associated with any measurement: accuracy and precision. The first one requires the minimization of PCA reconstruction bias; the second topic deals with the potential improvement of uncertainty interval due to PCA capacity of using the redundancy of the whole dataset. We will tackle these two different points using both experiments and simulations, and the results will be described in a series of two papers (Part I and Part II respectively). For EDS studies, the main contribution arises from counting noise (leading to PCA random noise bias), and the present manuscript reports a thorough analysis of the minimal level that the SNR must exceed to guarantee that PCA reconstruction carries the information of true NP composition variation in the sample. We have analyzed the applicability of several suggested PCA bias estimators (Malinowski, 2002; Faber et al., 1995a, 1995b; Nadler, 2008, 2009) and suggested some criteria to predict the reliability of derived chemical composition distributions.

2. Materials and methods

Nanoparticle synthesis. Bimetallic (Au_xAg_{1-x}) NPs have been produced using a homemade gas aggregation cluster source where a cylindrical magnetron is used to sputter atoms from a central target made of twisted Au and Ag wires (de Sá et al., 2014). A series of electrostatic lenses are used to reduce NP kinetic energy to achieve a "soft landing" on the TEM grid (for a ~4 nm in diam. NP, this corresponds to ~0.05 eV/atom). The NP size distribution contained in the molecular beam can be followed in-situ by

time-of-flight mass spectrometer (TOF-MS). TEM images indicate that NP size distribution follows a log-normal function (mean diameter ~4 nm, width ~3 nm in width, see Figure S1). Three different bimetallic nanoparticle samples (A, B, C) have been used in the present study whose Au content in atomic fraction is (0.73±0.01), (0.55±0.02) and (0.48±0.01) respectively. The chemical composition has been measured on an ensemble of NPs using an open parallel TEM beam with a long counting time to increase x-ray total intensity while minimizing radiation damage (dose ~ 10 $e^-/Å^2$).

Electron microscopy: data acquisition and processing. We have used several STEM microscopes to compare and optimize experimental configurations and beam energy values: JEM-2100F (SDD 30 mm^2, LNNANO-Campinas-Brazil), JEM-F200 (two SDD, 0.8 sR, JEOL), FEI-Tecnai G20F (Si(Li) 30 mm^2, LCE-UFSCAR-Brazil), TitanThemis (Super X Quad SDD, 0.8 sR, LNNANO-Campinas-Brazil). Microscopes were set up for ~ 0.5 nm pixel size, probe diameter ~ 0.7-1.0 nm, and a dwell time of ~200 ms per pixel for the sake of comparison between instruments. We have used different TEM instruments and, for all experiments, specimens were mounted in a low background Be holder. On average, each HSI image contains about 5-8 particles, then several HSIs are necessary to gather statistical meaningful description of the NP sample.

Measured EDS spectra have been binned to get a total of 512 energy channels, in order to increase SNR for PCA processing. Automatic recognition of NPs on the chemical maps (binarization and segmentation) has been performed using Scikit-image Python library (van der Walt et al., 2014). We have calculated the total x-ray characteristic signal (Au-Mα and Ag-Lα peaks) from individual particles by adding signal from the pixels located inside each NP region, because it is well known that addition maintains the Poisson nature of the sum.

The quantitative analysis of the NP chemical composition followed the Cliff-Lorimer approach

(Eq. 1, C_x *and* I_x are the atomic fraction and the x-ray intensity of element x respectively, (Cliff &Lorimer, 1975)). The Cliff-Lorimer factors (k_{AB}) were derived experimentally using a thin film of known composition ($k_{AuAg} = 0.93 \pm 0.01$ for data from Experiment I and $K_{AuAg} = 0.85 \pm 0.01$ for Experiment 3, simulations used $k_{AgAu} = 1$). The film was generated using a multilayer thin film calibrated using a quartz balance to control the atomic ratio through the relative film thickness.

$$\frac{C_{Au}}{C_{Ag}} = k_{AuAg}\frac{I_{Au}}{I_{Ag}} \qquad \text{(Eq. 1)}$$

The calculation of chemical composition error bars has taken into account all EDS quantitative analysis steps applied on a fitted curved on the experimental EDS spectra (background removal, x-ray peak integration and composition determination). All EDS and HSI PCA processing steps have been performed using the open-source Hyperspy Python library (de la Peña et al., 2017).

We have performed experiments at 80 kV to increase electron ionization cross-section considering relativistic estimations (Zaluzec, 2019). It is important to consider that recent microscopes (as the AC microscope used here) show very good performance at low voltages. However, in practice, the reduction from the nominal microscope voltage (300/200 keV) to lower values (100-80 keV) may reduce electron gun efficiency and increase the effect of TEM optical aberrations. Consequently, the probe current may decrease significantly, producing an unwanted consequence of, effectively, reducing EDS counts. It is important to verify properly the microscope performance at a lower voltage, because the optimal setup is strongly dependent on the particular TEM instrument. In addition, lower voltages (80-100 keV) should be preferred for lower atomic number samples in order to reduce knock-on damage (preferential damage mechanism for lower atomic number samples (Egerton et al., 2010; Egerton, 2012, 2019; Braidy et al.,

2008). We would like to ~~remind~~ note that all experiments discussed in this study have been acquired using the same dwell time (200 ms) for the sake of comparison between different experimental conditions.

Our experiments accumulated several frames for the final HSI (Jones et al., 2018), what minimizes the dose rate and extends the sample useful life. Instead of taking a single image with dwell time 200 ms, we have taken a series of images of the same region (for example 10 scans at 20 ms dwell time), and subsequently, added the individual frames. An important issue related to alloy NPs analysis is guarantying that composition is not modified along measurement time by radiation damage. As Ag atoms should be sputtered more easily than the Au ones, we would expect a gold enrichment with time (Egerton, 2019). To analyze this issue, we have taken a long series of 20 ms dwell time scans and followed the chemical composition evolution. This allowed us to verify that no detectable composition changes occur for the chosen dwell time in this work (200 ms).

At present, high spatial resolution chemical mapping is routinely performed with good reliability and reasonably short acquisition times, but spectra from individual pixels typically have insufficient counts for proper statistical quantification. According to the basic procedures of EDS microanalysis, a characteristic peak occurrence can be identified if the signal is three times (3x) the background noise (σ_{BGD}). This criterion, however, only applies to detection, and is inadequate for quantitative EDS chemical analysis. The seminal work of Currie (Currie, 1968, 1999; Belter et al., 2014; Williams & Carter, 2016)) shows that a quantitative analysis requires a significant increase of the signal level (approximately 15 timed higher than the uncertainty interval) to attain the measurement of a quantity with 10% relative error. This increases enormously the signal required, in particular when the detection suffers from shot noise, given that in Poisson statistics intensity uncertainty is the square root of the intensity itself, $\sigma_I = \sqrt{I}$.

PCA Processing. PCA takes as basic input a 2D matrix D of dimension mxn; m represents the number of pixels that have been acquired and, n the number of channels in the spectra (Jolliffe, 2002). This matrix description shows that PCA does not consider any specific information about spatial location of the pixels in the experimental image. When a spectral correlation exists between pixels it will show up as a significant weight (score) for a particular PCA component (loading).

PCA processing provides a number of orthogonal eigenvectors encoding the information contained in data variance. The eigenvalues (or singular values) associated with each component express their information content and are usually displayed in the so-called scree-plot (Jolliffe, 2002; Jolliffe & Cadima, 2016; Titchmarsh, 1999). In our experiments, the scree-plots show a clear and regular profile with a well-defined kink, where the number of components showing data variance visibly higher than noise can be easily identified (see example in Figure 1).

The principal source of noise in an EDS HSI is counting noise of Poisson nature; then before PCA processing, we have performed the variance stabilization using a data scaling procedure (Keenan & Kotula, 2004; Kotula & Keenan, 2006; Kotula & Van Benthem, 2015). The scaled data set matrix D_s has been calculated from $D_s = GDH$, where G is the spatial scaling factor, D is the raw data, and H is the spectral scaling factor (as implemented in Hyperspy (de la Peña et al., 2017)). Recently, it has been ~~have~~ suggested that for a very sparse data matrix obtained from NP samples (Kotula & Van Benthem, 2015), better multi-variate decomposition results are obtained if the scaling only considers the spectral scaling factor ($D_s = DH$). These authors analyzed a rather high sparse measurement (99.8 %). Our experiments have been acquired with a pixel area of 0.5x0.5 nm^2=0.25 nm^2, considering a particle of 6 nm in diameter, its area is about 28 nm^2, which should be covered by approximatively 110 pixels. This simple calculation shows that the sparsity level of in about 85-90% (scans of 64x64 pixels containing about 5-8 NPs, about 80-90%, [1-(660/4096)]~0.84 or 84%).). As the sparsity of our data is not so extremely low, we have

used the conventional complete spatial and spectral scaling in this work; the possible influence of changing the scaling procedure will be analyzed in future work.

In this report, the analyzed samples are very close to ideal ones: binary alloy particles (with a certain composition variation) distributed on a uniform thin substrate (a-C). As a consequence, the information content of each PCA principal component can be understood reasonably well; simulations of high SNR HSI indicate three dominant loadings without significant information mixing (see Figure 2): a) 1st component associated to TEM support grid and another EDS spurious x-ray sources (Fe from magnetic lenses, Cu from grid, etc.), what is due to the fact that in our experiments and simulations most of pixels (~85%) do not contain any NP EDS signal; b) 2nd component containing information on the mean NP EDS spectrum and mean chemical composition defined by the Au and Ag counts ratio of this component ($I_{Au,PC\#2}/I_{Ag,PC\#2}$); and, finally, c) 3rd component accounting for variation of composition among different particles (Potapov & Lubk, 2019). We must note that the 3rd component represents an unphysical spectrum (Potapov, 2016), because the EDS peaks (Au-M and Ag-L edges) show up with opposite signs for the two elements (positive/negative, see Figures 2c), as mathematically required to induce composition variation (or different I_{Au}/I_{Ag} ratio values).

As a rule of thumb, some components beyond the scree-plot elbow (2-3 additional components) are included in the reconstruction in order to include some remaining information carried by the first components associated with noise. All reconstructions in this work have included 5 components to render easy comparison between different experiments and simulations of binary alloy NPs (see Figure 1).

Simulated Datasets. Simulations offer irrefutable advantage against an experiment because the number, nature and strength of latent factors can be fully controlled and the related effects followed in detail. Also,

as the "true" object is precisely known, it is possible to compare the same object under situations of different noise levels.

It is critical that model spectra reproduce key features associated with measurements (background, energy calibration and magnitude of the typical signals) which strongly influence PCA output. We must be aware that most of simulations do not include complex factors such as plural scattering (critical issue for EELS) or EDS sum/escape peaks, etc. In fact, several of these issues are consciously neglected because they might severely complicate interpretation between information content of the data set and the PCA output).

Our simulations consider binary alloy NPs on a TEM grid and they include all other elemental EDS peaks related to the Cu grid and sample environment (Fe, Ni, etc.); the intensities of these peaks were scaled in relation to Au and Ag ones in agreement with our experimental data. After a noise-less spectra is calculated, random Poisson noise was added to data in order to yield a realistic model of experiments.

Once the total counts per NP to be considered in the simulation is defined, these counts are distributed inside pixels forming a circle (NP projection) in the simulated HIS. It is important to emphasize that counts distribution inside the NP associated pixels considers the spherical shape of the NPs, such that counts and SNR diminish gradually from pixels located from center to surface (thickness variation along electron path). As a final step, we add the random Poisson noise to the voxels x-ray intensity. This issue is essential for a faithful description of experiments, as it induces variation in the overall EDS SNR for each NP pixel, which will strongly influence the PCA output.

3. Results

Figure 3 shows an example of a typical EDS HSI data from AuAg nanoparticles taken in an uncorrected 200 kV FEG TEM/STEM (described in the text as Experiment I); the spectrum in Figure 2a

represents a good illustration of counting level that can be achieved. On average, we observe I_{Au}~*380* total counts for a ~6 nm NP, resulting in an intensity relative error of approximately $\Delta I/I = 5\%$ considering Poisson noise ($\sigma_I = \sqrt{I}$~ 19 cts).

For these experiments, we have decided to apply the PCA processing to each HSI individually, because, in many cases, the HSIs were acquired several weeks apart. Grouping several HSIs may be useful to exploit additional redundancy, but noise components may vary between images. For example, a thicker substrate or the change of grid contribution due to a scan located closer to a grid bar may induce the appearance of additional meaningless principal components. These issues were not significant in our experiments (usually a square scan, ~30 nm side), but a careful analysis of EDS intensity showed that the beam current of the used microscope (more than 10 years old) was varying significantly. Then each scan showed a slightly different SNR, which hindered their grouping into a big data set and, analyzed by a single PCA processing step. For data acquired with a modern AC TEM/STEM stability and reproducibility were not an issue and the dataset could be built by the stacking of several (64x64 pixels) HSI .

After PCA reconstruction, the automatic numerical recognition of smaller NPs on the chemical maps has been more efficient; this is even more pronounced for very small particles, 2-3 nm in diameter (Figures 3e,f). PCA treatment generates a scree-plot where two components display much higher variance than noise (see Figure 1); as expected, these two loadings carry information about the sample support and the average NP spectrum, respectively. In figure 4a, raw and PCA derived composition estimations (C_{Au}) of single NPs are distributed around the expected mean composition value. The analysis of composition distributions histograms at the right side of Figure 4a yields a standard deviation (STD) of $\Sigma_{C_{Au}den}$~0.022 vs. $\Sigma_{C_{Au}raw}$~0.05, (hereafter symbol Σ will be used to describe the calculated STD of a distribution ${\Sigma_{STD}}^2 = \sum(x-\bar{x})^2/(n-1)$). The width of calculated concentration distribution obtained after PCA denoising is much narrower than derived from raw data. This comparison indicates a strongly biased PCA

reconstruction: due to the rather low SNR, PCA processing has been unable to rank properly the information on intensity variations between NPs (expected to be contained in the 3rd principal component, Figure 2c, (Potapov & Lubk, 2019)). Then, the ratio (I_{Au}/I_{Ag}) and the chemical composition derived from it is almost completely fixed and determined by the 2nd principal component (representing the average NP spectrum). This generates a strong correlation between I_{Au} and I_{Ag} intensities significantly narrows the chemical composition distribution calculated for the ensemble of NPs. It is important to note that PCA-treated points are displayed without error bars, because there is no accepted and well-defined procedure for calculating confidence intervals after denoising (this issue will be addressed in a forthcoming paper associated with Part II of this work (Moreira, 2021)).

To test the above interpretation, we have built a dataset containing a bimodal distribution of chemical compositions by stacking a series of 4 HSI measured with the same microscope on two different NP samples (A-2x, B-2x, C_{Au}~0.50 and ~0.75 Au respectively, hereafter noted Experiment II). Using raw data, chemical analysis can distinguish the separate composition distributions (see Figure 4b, the calculated composition uncertainty bar was estimated to be $\sigma_{C_{Au}}$~ 0.04 or 4 at%). After PCA treatment, the chemical composition distribution from Exp. II appears significantly different from raw data: it is quite narrow and centred around the average composition of the two NP populations. The results derived from Exp. I and Exp. II, indicate that this SNR level merely allows the reliable determination of *average* ensemble chemical composition, and information on individual NP intensity remains indistinguishable from noise (variations of I_{Au}/I_{Ag} should carried by the 3rd principal component for a high SNR dataset, see Figure). This represents a clear example of *random noise* PCA bias, showing that insufficient counting may yield unrealistic results. If this PCA reconstructed dataset is analyzed with hubris, serious misinterpretations may follow about the chemical homogeneity of the analyzed sample.

The logical solution to make further progress is to increase the SNR by collecting more x-ray counts. Reducing the incident beam energy from 200 to 80 keV increases the ionization cross-section (and consequently the EDS signal) by a factor of 1.6x considering relativistic estimations (Zaluzec, 2019). We have performed experiments at 80 keV using an aberration-corrected (AC) STEM microscope equipped with four EDS detectors (named Experiment III). A significant counting improvement has been attained, leading to a total integrated count of $I_{Au} = 1950$ counts for a 6 nm NP (see Figure 5). Considering noise, the obtained ~5-fold gain in counts, just leads to a 2-fold decrease of intensity relative error $\Delta I/I$ ~2% ; which clearly exemplifies how difficult and complex is to increase SNR in data subjected to Poisson statistics. For this experiment, the raw dataset was built by grouping 4x acquisitions of 64x64 HIS. PCA treatment of Experiment III shows a substantial improvement of the reconstruction; Figure 5c shows ~~a reasonable~~ reasonably good agreement between chemical composition derived from raw and denoised data, indicating a strong reduction of *random noise* bias. The majority of measured NPs show similar composition within experimental confidence level (the measured composition distribution show similar widths as evidenced by the calculated standard deviations $\Sigma_{C_{Au,raw}} = 0.043$ and $\Sigma_{C_{Au,den}} = 0.031$). This amelioration is mainly associated to the better retrievability of a 3rd PCA component carrying information on individual NP intensity variations due to chemical composition variability (note the slight improvement of the 3rd component ranking in Figure 1b). Only few NPs display significant differences with raw data, suggesting the occurrence of some biasing; in fact, these points correspond to the smallest detected NPs (i.e. with smaller count value per NP, ex. I_{Au} ~ 400 cts, diameter ~3 nm).

4. Discussion

Experiment III provides evidence that it is essential to obtain measurements with counting level (or SNR) exceeding a certain minimal value to allow the minimization of *random noise* bias when analyzing alloy NP samples displaying a chemical composition variation. This allows the reliable ranking

of physically significant PCA components and produces good fidelity denoised PCA reconstruction (Figure 5c).

Comparison of Experiments I and III shows that PCA may allow reliable denoising, but it can not recover information that is not actually carried by the combination of data SNR and redundancy (data set size); this property has been named "EDS retrievability limit" (Potapov & Lubk, 2019). The observed bias is associated to inefficient retrieval of the "true" NP intensity variations, an information carried by the 3rd principal component (see Figures 1 and 2).

To advance further, it is important to develop criteria/expressions to gather predictions of potential PCA random noise biasing effects directly from the measured data (counts value, data matrix size, etc.). Looking carefully at our results, we may note that the comparison of composition histograms of raw and denoised data reveals the different quality of data reconstructions from Exp. I and Exp. III (Figures 4a and 5c), Then, it is interesting to suggest that the ratio of calculated composition distribution STDs or widths ($R_w = \Sigma_{C_{Au},den} / \Sigma_{C_{Au},raw}$) represents a good assessment of reconstruction fidelity and a potential indicator of occurrence of random noise biasing.

Many studies have addressed the effect of random measurement noise on PCA bias (Malinowski, 2002; Faber et al., 1995a, 1995b; Nadler, 2008, 2009) when noise is assumed to be uncorrelated and homoscedastic (all its random variables have the same finite variance, σ_H^2). In contrast, our experiments involve Poisson noise, where the variance is dependent on the value of the particular voxel intensity ($\sigma^2=I$). Then, the straightforward application of several bias indicators must be performed with extreme care to predict potential biasing (i.e. displacement from "true solutions"). Additional complications to estimate random noise variance arise from the high sparsity of EDS HSI of NPs (Keenan & Kotula, 2004) (many spectral energy channels include zero or just a few counts). For example, only 10-15% of pixels in our 64x64 scans carry x-ray counts from the particles. This leads to an average counting value per voxel (

$I_{ave,voxel}$) that may be <<1 by several orders of magnitude. Then, an estimation of data Poisson noise from mean voxel intensity ($\sigma = \sqrt{I_{ave,voxel}}$) is nonfunctional, because noise becomes always larger than the measurement itself. From another angle, the variance stabilization through data scaling of EDS spectra before PCA processing (Keenan & Kotula, 2004; Kotula & Keenan, 2006; Kotula & Van Benthem, 2015) may induce additional significant error if sparsity is present.

Several studies (Cueva et al., 2012; Lichtert &Verbeeck, 2013; Spiegelberg & Rusz, 2017; Potapov, 2017; Verbeeck & Van Aert, 2004) have analyzed PCA eigenvalues bias using derivations of Faber et al. (Faber et al., 1995a, 1995b), where the eigenvalue shift contains two terms. The first one represents an additive contribution directly related to random noise variance, and the second term depends on the number of pixels *m* and spectral channels *n*. Furthermore, combining Fisher's information concept and the Cramérs-Rao inequality, Verbeeck et al (Verbeeck & Van Aert, 2004, Lichtert &Verbeeck, 2013) have been able to estimate a lower bound for systematical errors occurring in PCA reconstructed EELS data. This estimation requires the evaluation of a homoscedastic-like variance from the average count of the HSI voxels; as mentioned above, the sparsity of EDS HSIs strongly limits the applicability of these approaches.

In this work, we have explored the bias estimator derived from the so-called *spiked covariance model* (Nadler, 2008, 2009). This author has revealed the existence of an abrupt information-loss phase transition as a function of noise variance, leading to a well-defined bias estimator (E_{IL}, Eq. 6); a bias-free PCA reconstruction is always achieved when the estimator is less that a threshold value (T_{IL}). This bias estimator displays a simple mathematical expression describing the interaction between data set size (*n,m*) and the ratio of experimental noise (variance σ^2) and *noise-free data variance* (v^{*2}, hereafter * will indicate that the variable is the *true noise-free data value*). It is important to emphasize that *"true"* parameters are seldom accessible to the experimentalist, so the elaboration of predictions is somewhat challenging. In the

limit n,m tending to infinity, the threshold becomes $T_{IL} = 1$ and the criterion becomes $E_{IL} < 1$. The behavior of estimator has already been explored both in EELS and EDS HSI processing (Lichtert & Verbeeck, 2013; Potapov, 2016; Potapov & Lubk, 2019).

$$E_{IL} = \frac{n}{m}\left(\frac{\sigma^2}{\nu^{*2}}\right)^2 \qquad (2)$$

For our EDS experiments, the calculation of the E_{IL} using the average count per voxel yields very high values (10^7-10^{10}). It is essential to keep in mind that the interpretation of bias estimators must be analyzed in the context of the particular experimental constrains and targeted scientific problem. The EDS determination of NP elemental compositions requires the total number of counts per particle (I_{NP}), instead of the value of the individual spectral channels (or voxels) themselves. It is, therefore, crucial to identify a procedure to correctly evaluate the ratio of counting noise and *true data* variances (σ^2/ν^{*2}) considering how it affects the information we are looking for (assessment of composition variation between NPs), overcoming issues introduced by the intrinsic sparsity of the data.

For example, let´s consider a NP generating a Au-M line total counts of $I_{Au,NP}$; we will assume that these counts are homogeneously distributed over the number of pixels (m_{NP}) inside the particle region (for a scanning step 0.5 nm, this correspond to ~110 pixels for a 6 nm NP). Then, we distribute the estimated pixel counts (I_{NP}/m_{NP}) homogeneously on the spectral channels (n_{Peak}) under the EDS characteristic peak (for EDS peak ~200 eV wide and a spectrum channel width of 20 eV, we obtain n_{Peak}= 10). The Poisson noise variance for voxels under the EDS peak is ${\sigma_{I_{Au,vox}}}^2 = I_{Au,NP}/(m_{NP}n_{PEAK})$. In fact, even taking this simple estimation of voxel intensity, it yields a value < 1 for Exp. I, indicating strong

sparsity. Only for the count values in Exp. III, the average total cts/voxel results slightly > 1 and the counting error Poisson statistic (square root) can now be applied, always resulting in noise < measurement.

An essential parameter to evaluate the E_{IL} bias estimator is the noiseless data variance, so we must estimate how chemical composition dispersion influences the *noise-free* voxel x-ray count values. Let´s start with a rather simple model case, where the NP ensemble displays a Gaussian distribution of chemical composition (mean $C_{Au,0}$, variance $\Sigma^2{}_{C_{Au}}$). To go further, we must deduce how to convert the Au composition standard deviation $\Sigma_{C_{Au}}$ into a characteristic EDS intensity variation or number of Au counts ($\Delta I_{Au,NP}$).

Chemical compositions are derived from EDS intensity using the Cliff Lorimer expression in (Eq. 1) (Cliff &Lorimer, 1975). Considering the normalization of a binary AuAg alloy compositions ($C_{Au} + C_{Ag} = 1$), we can relate C_{Au} to the EDS signal from Au atom:

$$C_{Au}\left(1 + \frac{I_{Ag}}{k_{AuAg} I_{Au}}\right) = 1 \quad (\text{Eq. } 3)$$

As our experimental values for the Cliff-Lorimer factor have ~~resulted~~ been determined to be very close to 1, we will assume that $k_{AuAg}=1$, or in other terms that Au and Ag generate the same quantity of characteristic photons per atom. This approximation renders the calculation simpler for a binary alloy; if we consider that all particles have the same number of atoms, a change in the number of atoms of one element, reduces (or increases) its characteristic EDS peak by the same amount of count that increase (or reduces) the other element peak. In other words, the total number of characteristic x-ray counts generated by the NP is constant $(I_{Tot,NP} = (I_{Au} + I_{Ag}) = cnst.$ For $k_{AuAg}=1$, Equation (3) may be rewritten as $C_{Au}(I_{Au} + I_{Ag}) = I_{Au}$ or $C_{Ag}(I_{Au} + I_{Ag}) = I_{Ag}$. Then, we can deduce that

$$I_{Au} = C_{Au}(I_{Tot,NP}) \quad \& \quad I_{Ag} = C_{Ag}(I_{Tot,NP}), \quad \text{(Eq. 4)}$$

Let´s assume a NP population of average chemical composition $C_{Au,0}$ (showing total counts $I_{Au,0}$ and ratio $(I_{Au,0}/I_{Ag,0})$ at this concentration) and a STD $\Sigma_{C_{Au}}$. We can derive how many counts ($\Delta I_{Au,\sigma}$) must be added to $I_{Au,0}$ in order to get a NP with a new concentration $C_{Au,\sigma} = C_{Au,0} + \Sigma_{C_{Au}}$. Due to the direct proportionality relation of Eq. 4, a composition change of $\Sigma_{C_{Au}}$ is induced from an intensity change of ΔI_{Au} $= I_{Tot,NP}\Sigma_{C_{Au}}$

Next, we will use these approximations using experimental values; for example, let´s take $C_{Au,0}$= 0.5 (50% Au atoms content) then $C_{Au,0}/C_{Ag,0} = 1$ (or $I_{Au,0}/I_{Ag,0} = 1$). A 6 nm diameter NP from Experiment I, should show a total EDS signal $(I_{Au} + I_{Ag}) = I_{Tot,NP} =$ 2 $I_{Au,0}$ = 760 cts. In our STEM HSI, a particle of 6 nm in diameter contains ~ 110 pixels. Then, each image pixel in the NP image must contain on average $I_{Au,pixel}$ ~380/110~3 cts/pixel. These counts are distributed in the EDS peak, which we can consider with ~200 eV at the base (or 10 energy channels, if ΔE=20 eV); then we finally arrive at $I_{Au,voxel}$ ~ 380/(110x10)~0.34 cts per voxel in average, considering all pixels and energy channels in the HSI region containing the NP. Finally, the Poisson noise of each voxel is obtained by taking the square root $\sigma_{voxel,Poisson} = \sqrt{I_{voxel}}$ ~0.6.

Assuming a *"true"* chemical composition distribution with standard deviation $\Sigma^*_{C_{Au}} = 0.04$, then $\Delta I_{Au,NP} = I_{Tot,NP}\sigma_{C_{Au}}$= 760 x 0.04 = 30 cts for a single nanoparticle. This value will be considered the *"true"* sample composition dispersion expressed in *"true" counts* variation per NP, $v^*_{I_{Au,NP}}$. Subsequently We distribute these counts between the NP pixels and energy channels under EDS peak to get "*true*" *noise-free* intensity variations in the sample $v^{\dagger}_{I_{Au,vox}} = \Delta I_{Au,NP}/(m_{NP}n_{PEAK})$ (as made above for NP intensity). This yields the "true" voxel intensity STD associated with the composition dispersion in the NP ensemble ($v^{\dagger}_{Au,vox}$~ $30/(110x10)$~ 0.029 cts).

These noise and true *noise-free* variances estimations ($\sigma^2_{I_{Au,vox}}$ and $\nu^{\dagger 2}_{I_{Au,vox}}$) yield a robust assessment of the effect of counting noise and how chemical composition dispersion influences the true noiseless data variance of EDS intensity, and they can be used to evaluate the information loss criterion and analyze PCA bias in our experimental studies. Hereafter, we will use these noise values in all evaluations of information loss bias estimators and redundancy parameters (m,n) will be derived from the HSI matrix size. Taking as basis our best experimental results, we will consider the dispersion measured from Exp. III (our best measurement) as the *"true"* distribution concentration variance ($\nu^*_{C_{Au}}$=$\Sigma_{C_{Au},Exp.III}$ = 0.043, Figure 5c).

Table I shows main parameters associated with the evaluation of counting levels (per NP and per voxel), ratio of raw and denoised compositions distribution widths (R_w), and E_{IL} bias estimator for our experiments. Although Experiment III shows a counting noise ($\sigma_{I_{Au,vox}}$~1.3) which is about 10 times the "true" standard deviation ($\nu^{\dagger}_{I_{Au,vox}}$) due to composition dispersion, PCA exploits redundancy and it is able to obtain a PCA reconstruction conserving reasonably well the raw composition distribution (see Fidelity column, R_w).

Table 1. Experimental counts and bias estimators evaluated for the Experiments I-III, considering 6-nm NP, and a *"true"* composition dispersion $\nu^*_{C_{Au}}$ = 0.043.

	$I_{Au,NP}$ (cts)	$I_{Au,vox}$ (cts)	Bias (Qualitative)	$\sigma_{I_{Au,vox}}$ (cts)	$\nu^{\dagger}_{I_{Au,vox}}$ (cts)	Fidelity (R_w)	Inform. Loss (E_{IL})
Experim. I	380	0.35	Strong	0.59	0.029	0.44	19100
Experim. II	293	0.37	Strong	0.61	0.079	0.20	1460
Experim. III	1950	1.77	Moderate/Low	1.33	0.15	0.73	180

It is very encouraging to see that denoising has preserved the small EDS signal variations necessary to describe concentration dispersion after PCA processing of Experiment III. Nonetheless, this contradicts the value of the information loss criterion value (E_{IL} = 180), which is much higher than the expected information loss threshold value $T_{IL} = 1$ (Nadler, 2008). In order to deepen our understanding of bias, we will analyze PCA reconstructions from a series of simulated HSI, where it is possible to consider a series of conditions in such a way that different SNR situations can be distinguished by comparing different NP *"true"* composition dispersions (variances $v_{C_{Au}}^{*\,2}$) vs. total x-ray counting.

Firstly, we will consider a bimetallic NP system with a Gaussian chemical composition distribution (noiseless composition standard deviation of $v_{C_{Au}}^{*} = 0.02$). Figure 6 compares the distribution of chemical compositions (C_{Au}) for simulated HSI using raw and PCA reconstructed data. These simulations show two distinct behaviors; a strong bias occurs in the first case (low fidelity reconstruction, $I_{Au,NP} = 1780$ cts, ~ Experiment III counts) where the calculated composition distributions is much narrower after denoising ($\Sigma_{C_{Au},den} \sim 0.0002$, see histogram at Figure 6a, right plot). The second case ($I_{Au,NP} = 7100$ cts) seems to yield a much better output, although some narrowing of the chemical composition distribution of the NP ensemble is still observable ($\Sigma_{C_{Au},den}=0.014$, Fig. 6b), indicating that this counting level is still insufficient to fully retrieve composition dispersion information. A very good recovery of the true composition distribution width ($\Sigma_{Au,den}=0.019$, Figure 6c requires a further increase of Au counts per NP (10660 cts, ~ 5-fold increase on Exp. III counts). It is interesting to note the scree plots from simulations in Fig. 6a and 6c (high and low biased reconstructions) do no apparently show significant differences (Figure 7a,b). For the strongly biased simulation displayed in Fig. 6a ($I_{Au,NP}$~1780 cts), the 3rd principal component is mostly flat in the region of Au and Ag EDS signal (Figure 7c), generating no information of intensity variation associated to composition dispersion. In contrast, the low bias PCA reconstruction (Figure 6c) shows that the high SNR ($I_{Au,NP}$~11160 cts) generates a 3rd components (displaying well defined Au and

Ag EDS peaks of opposite sign). It is surprising that this very good retrieval of the 3rd component, is associated with an singular value in the scree-plot which very close to the noise related components (see scree-plot in Figure 7b)

By increasing the noiseless composition distribution width to $v^*_{C_{Au}} = 0.036$ (Figures 6d,e), we can analyze bias occurrence as a function of SNR and true composition dispersion. As expected, it is possible to get a high-quality reconstruction with much less counts, due to the larger *"true" variance* of intensity associated to the much wider composition variability. In analogy to the precedent case, a counting level of $I_{Au,NP}$ = 1730 cts produces significantly low fidelity and biased reconstruction ($\Sigma_{C_{Au},den}=0.005$), but just a 2-fold increase of counts ($I_{Au,NP}$= 3540 cts) allows an excellent recovery of underlying "*true*" variability (i.e. accuracy) of NP composition distribution ($\Sigma_{C_{Au},den}= 0.035$). This total counts value is experimentally achievable, as it just represents less than a 2-fold increase of Exp. III counts.

Finally, we have also used simulations considering a bimodal true chemical composition distribution. We have considered two particle populations with composition averages around $C_{Au,1} = 0.75$ and $C_{Au,2} = 0.65$, each with $v^*_{C_{Au},STD}=0.02$ (Figure 8). A strong bias is observed for rather low counting ($I_{Au,NP} = 650$ *cts,* Figure 8a) and the existence of two populations remains undetected, the composition distribution is narrowly packed around the average elemental composition (equivalent to Experiment II). After increasing counting ($I_{Au,NP} = 1650$ cts), PCA reconstruction reveals the existence of two populations with a clear indication of certain concentration distribution for each one. A further 4-fold increase in counts ($I_{Au,NP} = 6600$ cts) yields a reliable chemical composition assessment after denoising (see Fig. 8c). These very good results are obtained at lower counting levels than before (Figure 6), because the sample exhibits a quite wide "effective true" dispersion for the sample considered as a whole ($v^*_{C_{Au},STD}=0.053,$ wrapping the bimodal distribution). Briefly, simulations indicate that PCA reconstructions may attain an

excellent accuracy for chemical composition distribution estimation (shown in Figures 6c, 6e and 8c), when the SNR exceeds a certain threshold value.

At this point, we must carefully look at the PCA calculations (ex. scree-plot) to see if it can gather additional information of bias occurrence that can guide the analysis of experimental studies. Our simulations suggest that even for a very light increase in the ranking of the 3rd principal component in relation to noise component seems to increase significantly the accuracy of reconstructions (see scree-plots in Figures 1b and 7d). As all PCA reconstructions have included 5 components, we will assume that the 5th principal component represents just noise (see scree-plot in Figures 1 and 7). Then, the ratio of singular values of 3rd and 5th components ($R_\lambda = \lambda_3/\lambda_5$) may represent an interesting quantitative guideline to identify if PCA processing has been able to rank correctly the 3rd component and detect chemical composition variability.

Table 2 displays the different bias estimators, discussed in this article, (R_w , R_λ and E_{IL}) calculated for our simulations in Figure 6. The cases displaying the low bias reconstructions (Figures 6c and 6e) are correctly identified by R_w and R_λ showing values around 0.9 and 1.14 respectively. Also, the information loss estimator E_{IL} show the lowest values for the best PCA reconstruction, however the numerical values (~400) is much higher that the theoretical threshold value =1 for an abrupt information loss event. Then, it is yet unclear if it can be used to predict the fidelity of NP composition assessment, even more considering that its evaluation requires the knowledge of noiseless data variance. For example, our estimation of E_{IL} from Exp. III yield a smaller value (E_{IL}~180) than for very high signal SNR simulations (E_{IL}~400); this is unexpected and it is probably due to our too optimistic guess of the *"true"* concentration variance (extracted from Figure 5c). Both R_w and R_λ are readily accessible from measurements and data treatment without previous knowledge of "true" composition distribution, then their very good correlation with bias minimization suggest that they may be very useful for practical applications.

Figure 9 shows a plot of R_λ vs. R_w for datasets (experiment and simulations) considered in this work. All cases displaying low bias are at the right top corner of the graph, allowing a direct assessment of PCA denoising performance. Then, for sample of binary alloy NPs with approximatively a Gaussian composition distribution, we suggest a criterion for obtaining a low bias PCA reconstructions: $R_\lambda \geq 1.15$ and $R_w \geq 0.9$; it is important to mention that both R_λ and R_w must exceed the minimal values.

If the sample composition distribution is more complex and measurements are rather noisy (ex. bimodal distribution as in Exp. II and simulations of Figure 7), the meaningful information may be distributed among more principal components. A strong random noise bias is immediately revealed by a low R_w; however, a proper understanding of information content carried by components may require additional simulations and a careful analysis of scree-plots and loading profiles before to define a different R_λ criterion.

Table 2. Bias estimators evaluated for the simulated HSIs shown in Figure 6. Different true concentration dispersion and counting levels are compared (ensemble of 54x 6-nm NPs, number spectral channels n=500; image size m=64x384=24576 pixels, scanning step 0.5 nm).

	$v_{C_{Au}}$	$I_{Au,NP}$ (cts)	Bias (Qualitative)	R_w	R_λ	Inform.Loss E_{IL}
Simul. 1 (Fig. 6a)	0.02	1780	Strong	0.01	1.043	15400
Simul. 2 (Fig. 6b)	0.02	7100	Moderate	0.70	1.035	970
Simul. 3 (Fig. 6c)	0.02	10660	Low	0.90	1.138	430
Simul. 4 (Fig. 6d)	0.036	1730	Strong	0.13	1.010	1550
Simul. 5 (Fig. 6e)	0.036	3540	Low	0.97	1.149	370

Finally, we would like to provide a tool to roughly estimate EDS counts generated by NPs of different sizes and incident energy and detector geometries (remember that we have always used 200 ms dwell time, beam current ~400-500 pA). We have compiled the values of EDS counts per NP as a function of particle diameter for different instruments (uncorrected 200 kV and AC corrected 80 kV experiments) versus NP diameter); the experimental points were fitted with a 3rd-order polynomial (see Figure S2). We expect that this plot may be exploited to llow a rough estimation of expected SNR for metal nanoparticle EDS experiments.

In this work, we have mainly addressed random noise bias, and not discussed *underfactoring bias* (reconstruction using an insufficient number of principal components) (Malinowski, 2002; Faber et al., 1995a, 1995b; Nadler, 2008, 2009), because the information content of each PCA principal component can be understood reasonably well for our sample. We would like to emphasize Potapov & Lubk have reported an interesting tool to identify if the PCA component contains information or just carry noise (Potapov & Lubk, 2019) by plotting the scores for successive principal components: noisy ones generate isotropic round clouds, but information-carrying ones show a structured anisotropic cloud (see example in Figure S3).

Conclusions:

Given the outstanding progress of TEMs instruments and associated detectors, it is essential to exploit their capabilities to target "quantitative" interpretation for image, diffraction, or spectroscopic data. PCA data processing is a very powerful tool to improve the SNR of hyperspectral images and extract all underlying information contained in the data. Our high spatial EDS experiments and simulations show that if the SNR exceeds a minimal threshold value, it is possible to avoid random noise biasing in PCA

reconstructions and, then get chemical composition distribution with very good fidelity of true values (i.e. accuracy). In fact, the experimental requirements represent a compromise between dataset size (redundancy), counting leval , and sample true composition variance.

We have suggested a few quantitative estimators that may provide information about random noise bias occurrence, and they can be derived for PCA processing output without knowledge of "true" noiseless sample properties. However, numerical simulations may be necessary to analyze the particular nanostructured system, in order to design properly data acquisition parameters and support the criteria used for the interpretation. It is important to emphasize that the physical behavior may vary dramatically between NP systems, so it is essential that simulations reproduce experiments as close as possible.

A natural question arises: how should one analyze measurements of bimodal distributions or three-element metal alloys? Which of the components will carry information on local differences? It is not easy to answer this question with a general applicable recipe for all kinds of samples. It is important to mention factors such as vibrations, scan noise or filament current oscillation may generate additional principal components in the scree-plot, and chemical information can become scattered into several components, so one must be prepared to carefully examine the scree-plots and information carried by principal component profiles.

ASSOCIATED CONTENT *S1

Supporting Information: Details on NP size distribution. Experimental intensities as a function of NP size for different TEM instruments. Identification of information content of principal components though scatter plots of scores. Figures S1-S3 (PDF)

AUTHOR INFORMATION

Corresponding Authors

Daniel Ugarte- *Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas-UNICAMP, 13083-859, Campinas - SP, Brazil; Email: dmugarte@ifi.unicamp.br*

Varlei Rodrigues- *Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas-UNICAMP, 13083-859, Campinas - SP, Brazil; Email: varlei@ifi.unicamp.br*

Authors

Murilo Moreira- *Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas-UNICAMP, 13083-859, Campinas - SP, Brazil*

Matthias Hillenkamp- Institute of Light and Matter, University Lyon, Université Claude Bernard Lyon 1, CNRS, UMR5306, F-69622 Villeurbanne, France

Luiz H. G. Tizei − Laboratoire de Physique des Solides, Université Paris-Saclay, CNRS, 91405 Orsay, France; orcid.org/0000-0003-3998-9912

Giorgio Divitini- Department of Materials Science and Metallurgy, University of Cambridge, , Cambridge CB3 0FS, United Kingdom

Caterina Ducati- Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB3 0FS, United Kingdom

Mônica A, Cotta- *Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas-UNICAMP, 13083-859, Campinas - SP, Brazil*

Notes

The authors declare no competing financial interest.

Acknowledgements

D. U. acknowledges financial support from the Brazilian Agencies FAPESP (No. 2014/01045-0), CNPq (402571/2016-9 & No. 306513/2017-0) and FAEPEX-UNICAMP. . V. R. thanks funding from FAPESP (2007/01722-9) and CNPq (555647/2006-4 & 577046/2008-0). M.M. thanks funding from CNPq (No. 162541/2018-0). Access to the FEG-TEM/STEM from the Brazilian Nanotechnology National Laboratory is acknowledged (LNNANO, grant No. ME - 22329). M. A. C. acknowledges financial support from FAPESP (Nos. 2013/02300-1 and 2013/10957-0) and CNPq (No. 479486/ 2012-3). L. H. G. T., G. D and C.D. acknowledge funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 312483 (ESTEEM2). M. H. acknowledges financial support from FAPESP (No. 2016/12807-4)

References

Alloyeau, D., Mottet, C., Ricolleau, C. (2012). *Nanoalloys: Synthesis, Structure and Properties*. London: Springer-Verlag.

Belter, M., Sajnóg, A., Barałkiewicz, D. (2014). Over a century of detection and quantification capabilities in analytical chemistry – Historical overview and trends. *Talanta* ***129***, 606-616.

Binns, C. (2014) *Nanomagnetism: Fundamentals and Applications*. *Vol. 6*. Amsterdam, Netherlands: Elsevier.

Braidy, N., Jakubek, Z.J., Simard, B., Botton, G.A. (2008). Quantitative Energy Dispersive X-ray Microanalysis of Electron Beam-Sensitive Alloyed Nanoparticles. *Microsc. Microanal.* ***14***, 166-175.

Brown, K.A., Brittman, S., Maccaferri, N., Jariwala, D., Celano, U. (2020). Machine Learning in Nanoscience: Big Data at Small Scales. *Nano Lett.* ***20***, 2−10.

Burke, M.G., Watanabe, M., Williams, D.B., Hyde, J.M. (2006). Quantitative characterization of nanoprecipitates in irradiated low-alloy steels: Advances in the application of FEG-STEM quantitative microanalysis to real materials. *J. Mater. Sci.* ***41***, 4512–4522.

Cliff, G., Lorimer, G.W. (1975). The quantitative analysis of thin specimens. *J.of Microsc.* ***103***, 203–207.

Cueva, P., Hovden, R., Mundy, J.A., Xin, H.L., Muller, D.A. (2012). Data processing for atomic resolution electron energy loss spectroscopy. *Microsc. Microanal.* ***18***, 667-675.

Currie, L.A. (1968). Limits for qualitative detection and quantitative determination. Application to radiochemistry. *Anal. Chem.* ***40***, 586–593.

Currie, L.A. (1999). Detection and quantification limits: origins and historical overview. *Anal. Chim. Acta* ***391***, 127-134.

de Sá., A.D.T., Oiko, V.T.A., di Domenicantonio, G., Rodrigues, V. (2014). New experimental setup for metallic clusters production based on hollow cylindrical magnetron sputtering. *J. Vac. Sci. Technol. B* ***32***, 061804.

Egerton, R.F., McLeod, R., Wang F., Malac, M. (2010). Basic questions related to electron-induced sputtering in the TEM. *Ultramicroscopy* **110**, 991–997.

Egerton, R.F. (2012). Mechanisms of radiation damage in beam-sensitive specimens, for TEM accelerating voltages between 10 and 300 kV. *Microsc. Res. Tech.* ***75***, 1550–1556.

Egerton, R.F. (2019). Radiation damage to organic and inorganic specimens in the TEM. *Micron* ***119***, 72–87.

Faber, N.M., Meinders, M.J., Geladi, P., Sjiistrijm, M., Buydens, L.M.C., Kateman, G. (1995a). Random error bias in principal component analysis. Part I. derivation of theoretical predictions. *Anal. Chim. Acta* ***304***, 257-271.

Faber, N.M., Meinders, M.J., Geladi, P., Sjiistrijm, M., Buydens, L.M.C., Kateman, G. (1995b). Random error bias in principal component analysis. Part II. Application of theoretical predictions to multivariate problems. *Anal. Chim. Acta* ***304***, 273-283.

Ferrando, R. (2016) *Structure and properties of nanoalloys in Frontiers of Nanoscience*. pp 2–337, Amsterdam: Elsevier.

Hawkes**,** P.; Spence, J. C. H (2019). *Springer Handbook of Microscopy*. Switzerland: Springer.

Heiz, U., Landman, U.E. (2007). *Nanocatalysis.* Heidelberg, Berlin: Springer.

Jolliffe, I. *T.* (2002). *Principal Component Analysis, 2 ed.* Springer-Verlag: New York.

Jolliffe, I.T., Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philos. Trans. A* ***374***, 20150202.

Jones, L., Varambhia, A., Beanland, R., Kepaptsoglou, D., Griffiths, I., Ishizuka, A., Azough, F., Freer, R., Ishizuka, K,, Cherns, D., Ramasse, Q.M., Lozano-Perez, S., Nellist, P.D. (2018). Managing dose-, damage- and data-rates in multi-frame spectrum-imaging. *Microscopy* **67**, 98–113.

Keenan, M.R., Kotula, P.G. (2004). Accounting for Poisson noise in the multivariate analysis of TOF-SIMS spectrum images. *Surf. Interface Anal.* ***36***, 203–212.

Kotula, P.G., Keenan, M.R. (2006). Application of multivariate statistical analysis to STEM Xray spectral images: interfacial analysis in microelectronics, *Microsc. Microanal.* ***12***, 538–544.

Kotula, P.G., Van Benthem, M.H. (2015). Revisiting noise scaling for multivariate statistical analysis. *Microsc. Microanal.* ***21***, 1423–1424.

Lichtert, S., Verbeeck, J. (2013). Statistical consequences of applying a PCA filter on EELS spectrum images. *Ultramicroscopy* ***125***, 35–42.

Lyman, C.E., Lakis, R.E., StengerJr., H.G. (1995). X-ray emission spectrometry of phase separation in PtRh nanoparticles for nitric oxide reduction. *Ultramicroscopy* ***58***, 25-34.

Malinowski, E.R. (2002). *Factor analysis in chemistry, 3rd ed*. New York, Wiley.

Moreira, M., Hillenkamp, M., Divitini, G., Tizei, L.H.G., Ducati, C., Cotta, M.A., Rodrigues, V., Ugarte, D. (2021). In preparation.

Mukherjee, P., Zhang, Y., Kramer, M.J., Lewis, L.H., Shield, J.E. (2012). $L1_0$ structure formation in slow-cooled Fe–Au nanoclusters. *Appl. Phys. Lett.* ***100***, 211911.

Mukherjee, P., Manchanda, P., Kumar, P., Zhou, L.. Kramer, M.J., Kashyap, A., Skomski, R., Sellmyer, D., Shield, J.E. (2014). Size-induced chemical and magnetic ordering in individual Fe–Au nanoparticles. *ACS Nano* ***8***, 8113–8120.

Nadler, B. (2008). Finite sample approximation results for principal component analysis: a matrix perturbation approach. *Ann. Stat.* ***36***, 2791–2817.

Nadler. B. (2009). Discussion, *J. Am. Stat. Assoc.* ***104***, 694-697.

Odom, T.W., Schatz, G.C. (2011). Introduction to plasmonics. *Chem. Rev.* ***111***, 3667–3668.

Parish, C.M., Brewer, L.N. (2010). Multivariate statistics applications in phase analysis of STEM-EDS spectrum images. *Ultramicroscopy* ***110***, 134–143.

Pennycook, S. J.; Nellist, P. D. (2011). *Scanning Transmission Electron Microscopy*. Springer: New York.

Potapov, P. (2016). Why principal component analysis of STEM spectrum images results in abstract, uninterpretable loadings? *Ultramicroscopy* ***160***, 197–212

Potapov, P. (2017). On the loss of information in PCA of spectrum-images. *Ultramicroscopy* ***182***, 191–194.

Potapov, P., Lubk, A. (2019). Optimal principal component analysis of STEM XEDS spectrum images. *Adv. Struct. Chem. Imag.* ***5***, 4.

Rossouw, D., Burdet, P., de la Peña, F., Ducati, C., Knappett, B.R., Wheatley, A.E.H., Midgley, P.A. (2015). Multicomponent Signal Unmixing from Nanoheterostructures: Overcoming the Traditional Challenges of Nanoscale X-ray Analysis via Machine Learning. *Nano Lett.* ***15***, 2716−2720.

Schlossmacher, P., Klenov, D.O., Freitag, B. (2010). Enhanced detection sensitivity with a new windowless XEDS system for AEM based on silicon drift detector technology. *Microsc. Today* ***18***, 14–20.

Spiegelberg, J., Rusz, J. (2017). Can we use PCA to detect small signals in noisy data? *Ultramicroscopy* ***172***, 40–46.

Thomas, J R., Leary, R.K., Eggeman, A.S., Midgley, P.A. (2015). The rapidly changing face of electron microscopy. *Chem. Phys. Lett*. ***631–632***, 103-113.

Titchmarsh, J.M., Dumbill, S. (1996). Multivariate statistical analysis of FEG STEM EDX spectra. *J. Microsc.* ***184***, 195–207.

Titchmarsh, J.M. (1999). EDX spectrum modelling and multivariate analysis of sub-nanometer segregation. *Micron* ***30***, 159–171.

van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart,E., Yu, T (2014). The scikit-image contributors scikit-image: image processing in Python. *PeerJ*. ***2***, e453.

de la Peña, F., Ostasevicius, T., Fauske, V.T., Burdet, P., Jokubauskas, P., Nord, M., Prestat, E., Sarahan, M., MacArthur, K.E., Johnstone, D.N., Taillon, J., Caron, J., Furnival, T., Eljarrat, A., Mazzucco, S., Migunov, V., Aarholt, T., Walls, M., Winkler, F., Martineau, B., Donval, G., Hoglund, E.R., Alxneit, I., Hjorth, I., Zagonel, L.F., Garmannslund, A., Gohlke, C., Iyengar, I., Chang, H.-W. (2017). hyperspy/hyperspy: HyperSpy 1.3 DOI:10.5281/ZENODO.583693.

Verbeeck, J, Van Aert, S. (2004) Model based quantification of EELS spectra. *Ultramicroscopy* ***101***, 207–224.

Watanabe, M., Kanno, M., Okunishi, E. (2010). Atomic resolution elemental mapping by EELS and XEDS in aberration corrected STEM. *Jeol News* ***45***, 8–12.

Williams, D.B.; Carter, C.B. (2009). *Transmission electron microscopy Part 1: Basics, 2nd ed.* Springer: New York.

Williams, D.B., Carter, C.B (2016). *Transmission Electron Microscopy: Diffraction, Imaging, and Spectroscopy*. Springer: New York.

Zaluzec, N.J. (2019). Improving the sensitivity of X-ray microanalysis in the analytical electron microscope. *Ultramicroscopy* ***203***, 163–169.

Figure Captions.

Figure 1. Comparison of scree-plots derived from EDS measurements on AuAg alloy nanoparticles experiments acquired using a non-corrected FEG TEM /STEM (a) and a modern aberration corrected microscope equipped with multiple EDS detectors (b). Reconstruction derived from experiments associated with (a) showed a strong random noise bias, while HSI described by (b) showed a good PCA reconstruction (see text for explanations). Note that the variance of the 3rd principal component is slightly above noise component in (b). Triangular points represent the components used for data reconstruction in this work.

Figure 2. Principal component profiles derived from a high SNR simulated HIS of bimetallic AuAg nanoparticles; for this ideal case, it is possible to make a clear identification of carried information. The loading in (b) includes Au and Ag peaks with the same signal (both negative, allowing the calculation of the NP average NP chemical composition); in contrast the component in (c) shows Au and Ag peaks with opposite signal (negative-positive) then it is possible to induce concentration variations in the PCA reconstruction.

Figure 3. Typical EDS HSI data from AuAg alloy nanoparticles (uncorrected FEG-STEM, 200KV, 64x64 pixels scan). Comparison of raw and denoised EDS spectrum from a ~6 nm NP for a single pixel (a) and the entire NP (b) where the orange curve represent the final fit of EDS processing derived from Hyperspy software used to integrate the peak intensity. c), and d) raw and denoised Au chemical maps, respectively; the corresponding NP identification is displayed in (e) and (f) respectively (NP pixels are presented in

different colors render easy the nanoparticles recognition). Note the increase in compactness of detected particles after PCA processing.

Figure 4. Quantitative chemical composition analysis of alloy NPs. a) Au atomic concentration derived from Exp. I using either raw or PCA denoised HIS (dashed line indicates average composition). b) Au atomic concentration from Exp. II, including a bimodal chemical composition distribution (atomic concentration C_{Au}~0.5 and ~0.75). Resulting chemical composition distributions are plotted at the right. The two populations are easily identified from raw data, while denoising generates a strongly biased results and NP composition is close to the average of the ensemble of NPs.

Figure 5. Typical EDS HSI data from AuAg alloy nanoparticles obtained for Experiment III (AC STEM, 4 SDD detectors, 80KV, 64x64 pixels scan). Comparison of raw and denoised EDS spectrum from a ~6 nm NP for a single pixel (a) and the entire NP (b) where the orange curve represent the final fit of EDS processing derived from HyperSpy software used to integrate the peak intensity. c) Comparison of chemical composition calculations derived using raw and denoised data; resulting chemical composition distribution are plotted at the right.

Figure 6. Effect of denoising on the assessment of chemical composition distribution deduced from simulated data sets including different *"true"* composition dispersion and counting levels. a-c) NP sample of average composition $C_{Au} = 0.75$ and distribution width $v^*_{C_{Au},STD}$=*0.02*. d-e) NP sample of average composition $C_{Au,1} = 0.75$ and a wider distribution $v^*_{C_{Au},STD}$=*0.036* (see text for explanations). The NPs are displayed following the increasing Au concentration along horizontal axis, allowing an easy

visualization of the composition distribution (histograms are displayed along the vertical direction on the right).

Figure 7. Scree-plot and 3rd principal component profile derived from the simulated HIS shown in Figures 6a and 6c. Note that for low counting level (1780 cts, (a)) the profile of principal component #3 (arrowed) does not carry any information of chemical composition variation (characteristic peaks from Au and Ag are not present (c)). In contrast, the low bias (or higher cts.) reconstruction of simulation from Fig. 6c is associated to a 3rd principal component with well formed Au and Ag peaks of opposite sign (d) in order to induce changes of the I_{Au}/I_{Ag} ratio that determines NP chemical composition.

Figure 8. Effect of denoising on the assessment of chemical composition distribution deduced from simulated data sets including bimodal "true" distributions of chemical composition (population 1: $C_{Au,1} = 0.75$, population 2: $C_{Au,2} = 0.65$, for both the distribution width is $v^*_{C_{Au},STD}$=*0.02*). The calculated chemical composition distribution is plotted at the right (see text for explanations).

Figure 9. Plot showing the correlation of proposed bias estimators R_w and R_λ for Experiments I and III and simulations displayed in Fig. 6 (index 1-5 indicated simulation describes in Fig. 6a-6f respectively).

Table 1. Experimental counts and bias estimators evaluated for the Experiments I-III, considering 6-nm NP, and a *"true"* composition dispersion $v^*_{C_{Au}} = 0.043$.

	$I_{Au,NP}$ (cts)	$I_{Au,vox}$ (cts)	Bias (Qualitative)	$\sigma_{I_{Au,vox}}$ (cts)	$v^{\dagger}_{I_{Au,vox}}$ (cts)	Fidelity (R_w)	Inform. Loss (E_{IL})
Experim. I	380	0.35	Strong	0.59	0.029	0.44	19100
Experim. II	293	0.37	Strong	0.61	0.079	0.20	1460
Experim. III	1950	1.77	Moderate/Low	1.33	0.15	0.73	180

Table 2. Bias estimators evaluated for the simulated HSIs shown in Figure 6. Different true concentration dispersion and counting levels are compared (ensemble of 54x 6-nm NPs, number spectral channels n=500; image size m=64x384=24576 pixels, scanning step 0.5 nm).

	$\nu_{C_{Au}}$	$I_{Au,NP}$ (cts)	Bias (Qualitative)	R_w	R_λ	Inform.Loss E_{IL}
Simul. 1 (Fig. 6a)	0.02	1780	Strong	0.01	1.043	15400
Simul. 2 (Fig. 6b)	0.02	7100	Moderate	0.70	1.035	970
Simul. 3 (Fig. 6c)	0.02	10660	Low	0.90	1.138	430
Simul. 4 (Fig. 6d)	0.036	1730	Strong	0.13	1.010	1550
Simul. 5 (Fig. 6e)	0.036	3540	Low	0.97	1.149	370

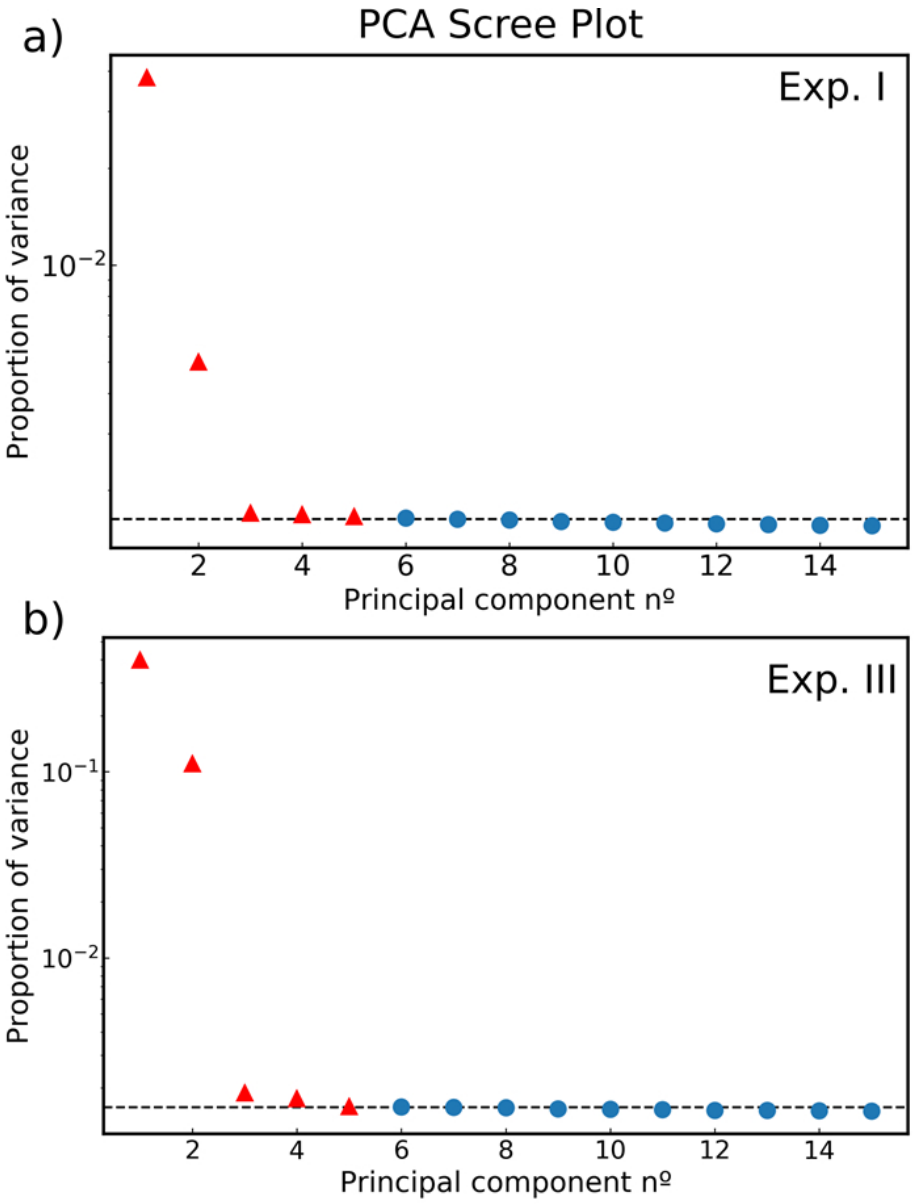

Figure 1. Comparison of scree-plots derived from EDS measurements on AuAg alloy nanoparticles experiments acquired using a non-corrected FEG TEM /STEM (a) and a modern aberration corrected microscope equipped with multiple EDS detectors (b). Reconstruction derived from experiments associated with (a) showed a strong random noise bias, while HSI described by (b) showed a good PCA reconstruction (see text for explanations). Note that the variance of the 3rd principal component is slightly above noise component in (b). Triangular points represent the components used for data reconstruction in this work.

248x322mm (72 x 72 DPI)

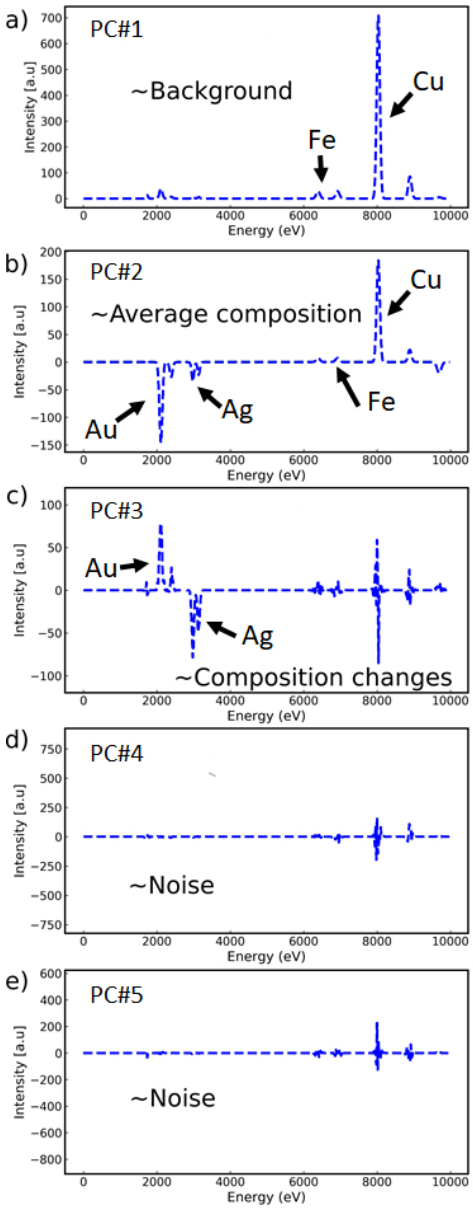

Figure 2. Principal component profiles derived from a high SNR simulated HIS of bimetallic AuAg nanoparticles; for this ideal case, it is possible to make a clear identification of carried information. The loading in (b) includes Au and Ag peaks with the same signal (both negative, allowing the calculation of the NP average NP chemical composition); in contrast the component in (c) shows Au and Ag peaks with opposite signal (negative-positive) then it is possible to induce concentration variations in the PCA reconstruction.

140x333mm (72 x 72 DPI)

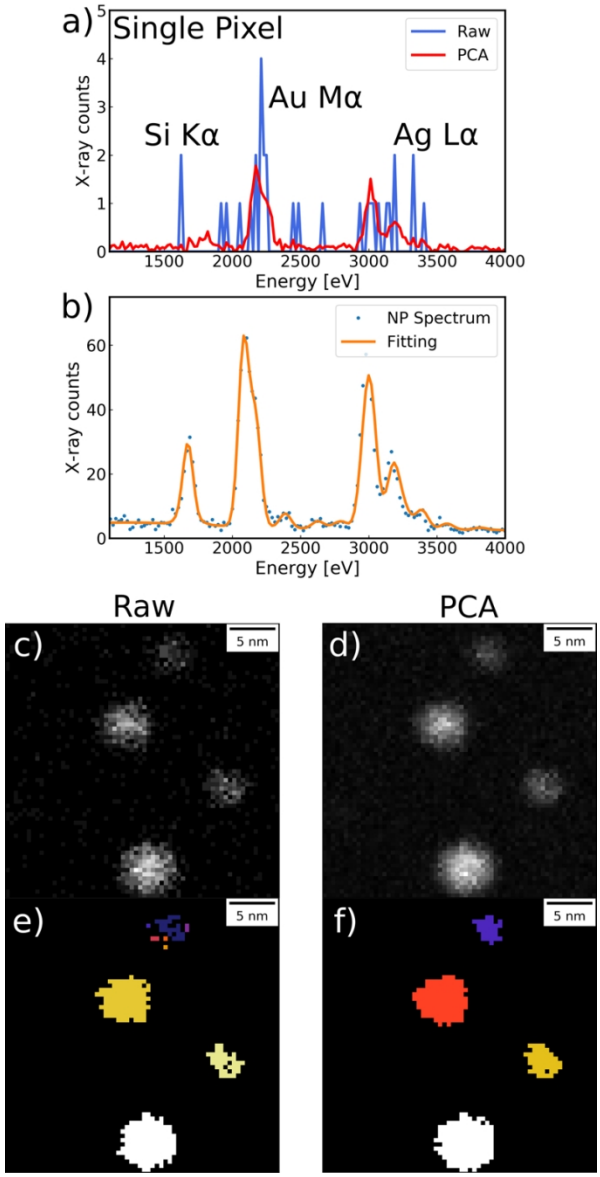

Figure 3. Typical EDS HSI data from AuAg alloy nanoparticles (uncorrected FEG-STEM, 200KV, 64x64 pixels scan). Comparison of raw and denoised EDS spectrum from a ~6 nm NP for a single pixel (a) and the entire NP (b) where the orange curve represent the final fit of EDS processing derived from Hyperspy software used to integrate the peak intensity. c), and d) raw and denoised Au chemical maps, respectively; the corresponding NP identification is displayed in (e) and (f) respectively (NP pixels are presented in different colors render easy the nanoparticles recognition). Note the increase in compactness of detected particles after PCA processing.

279x566mm (72 x 72 DPI)

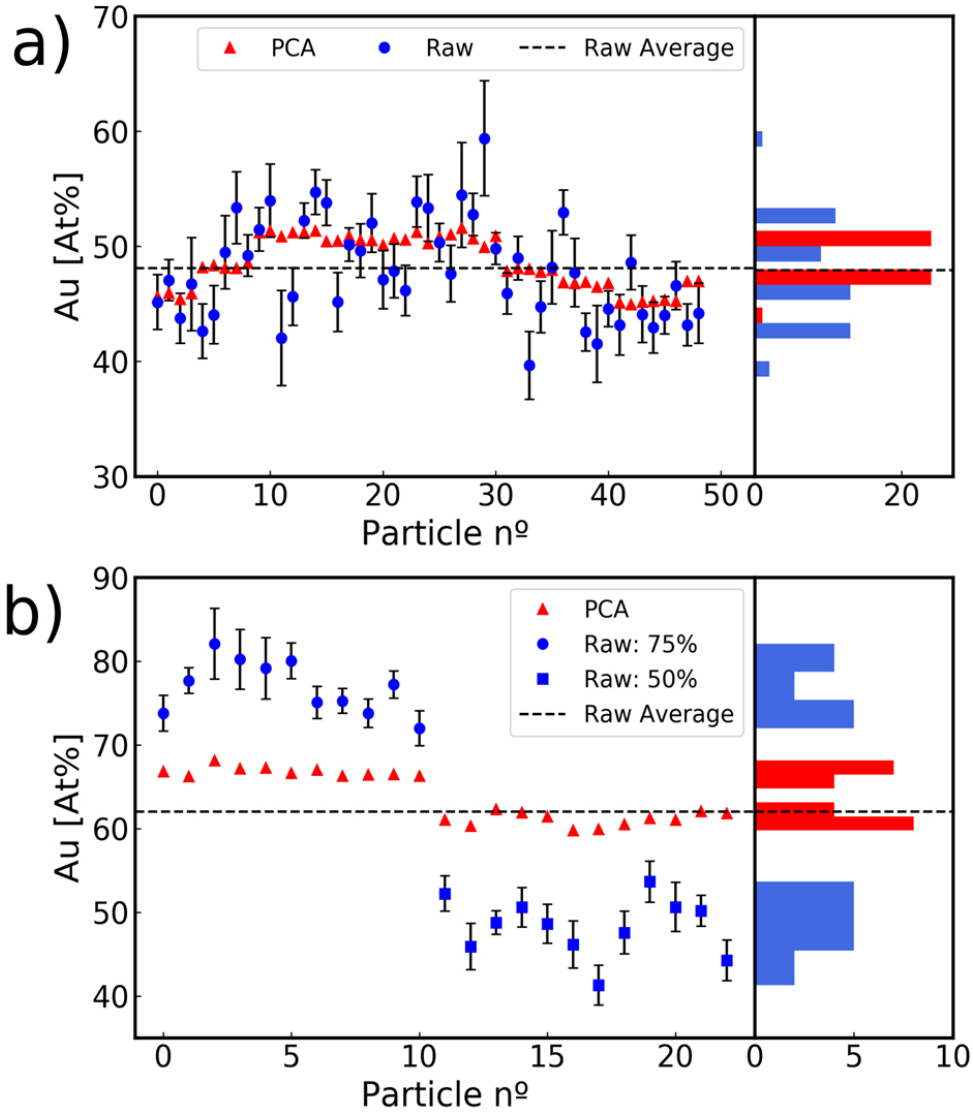

Figure 4. Quantitative chemical composition analysis of alloy NPs. a) Au atomic concentration derived from Exp. I using either raw or PCA denoised HIS (dashed line indicates average composition). b) Au atomic concentration from Exp. II, including a bimodal chemical composition distribution (atomic concentration CAu~0.5 and ~0.75). Resulting chemical composition distributions are plotted at the right. The two populations are easily identified from raw data, while denoising generates a strongly biased results and NP composition is close to the average of the ensemble of NPs.

341x399mm (72 x 72 DPI)

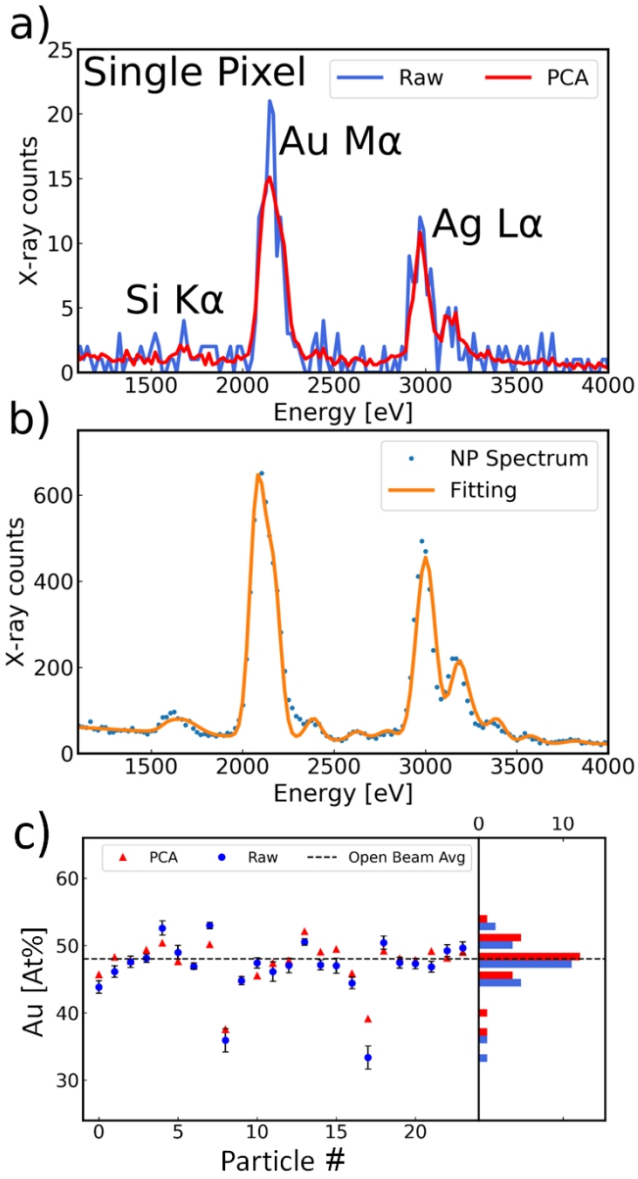

Figure 5. Typical EDS HSI data from AuAg alloy nanoparticles obtained for Experiment III (AC STEM, 4 SDD detectors, 80KV, 64x64 pixels scan). Comparison of raw and denoised EDS spectrum from a ~6 nm NP for a single pixel (a) and the entire NP (b) where the orange curve represent the final fit of EDS processing derived from HyperSpy software used to integrate the peak intensity. c) Comparison of chemical composition calculations derived using raw and denoised data; resulting chemical composition distribution are plotted at the right.

319x493mm (72 x 72 DPI)

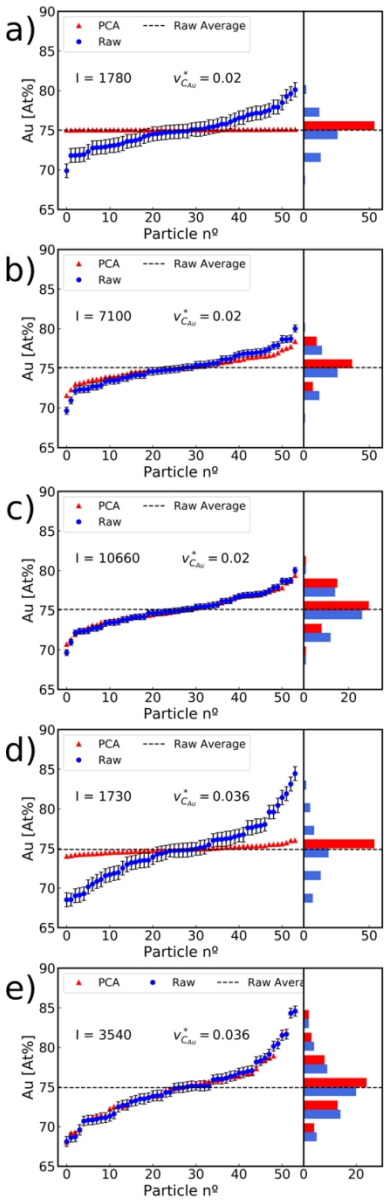

Figure 6. Effect of denoising on the assessment of chemical composition distribution deduced from simulated data sets including different "true" composition dispersion and counting levels. a-c) NP sample of average composition C_Au^ =0.75 and distribution width 〖v^*〗_(C_Au,STD)=0.02. d-e) NP sample of average composition C_(Au,1)^ =0.75 and a wider distribution 〖v^*〗_(C_Au,STD)=0.036 (see text for explanations). The NPs are displayed following the increasing Au concentration along horizontal axis, allowing an easy visualization of the composition distribution (histograms are displayed along the vertical direction on the right).

191x595mm (72 x 72 DPI)

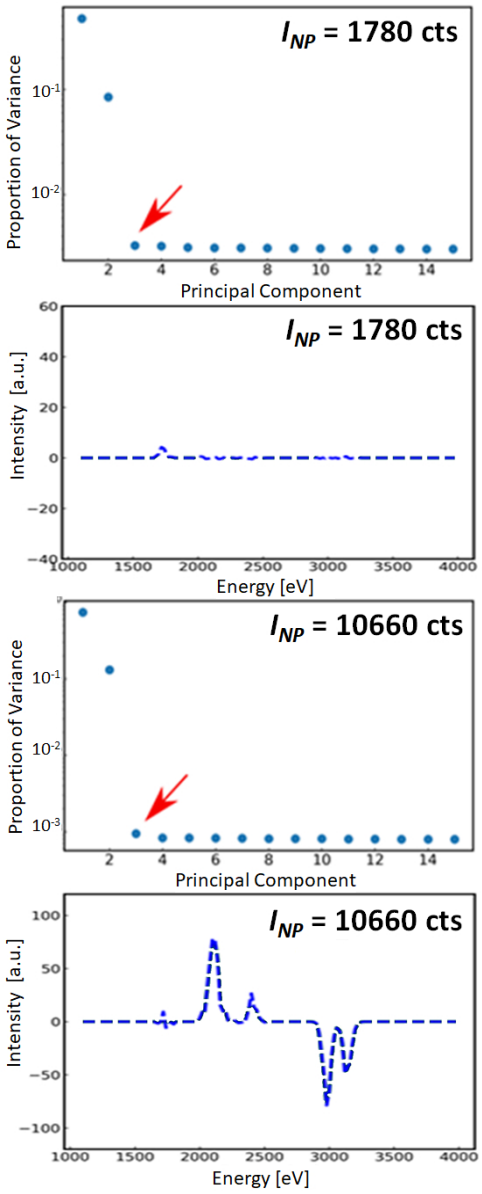

Figure 7. Scree-plot and 3rd principal component profile derived from the simulated HIS shown in Figures 6a and 6c. Note that for low counting level (1780 cts, (a)) the profile of principal component #3 (arrowed) does not carry any information of chemical composition variation (characteristic peaks from Au and Ag are not present (c)). In contrast, the low bias (or higher cts.) reconstruction of simulation from Fig. 6c is associated to a 3rd principal component with well formed Au and Ag peaks of opposite sign (d) in order to induce changes of the I_Au/I_Ag ratio that determines NP chemical composition.

200x481mm (72 x 72 DPI)

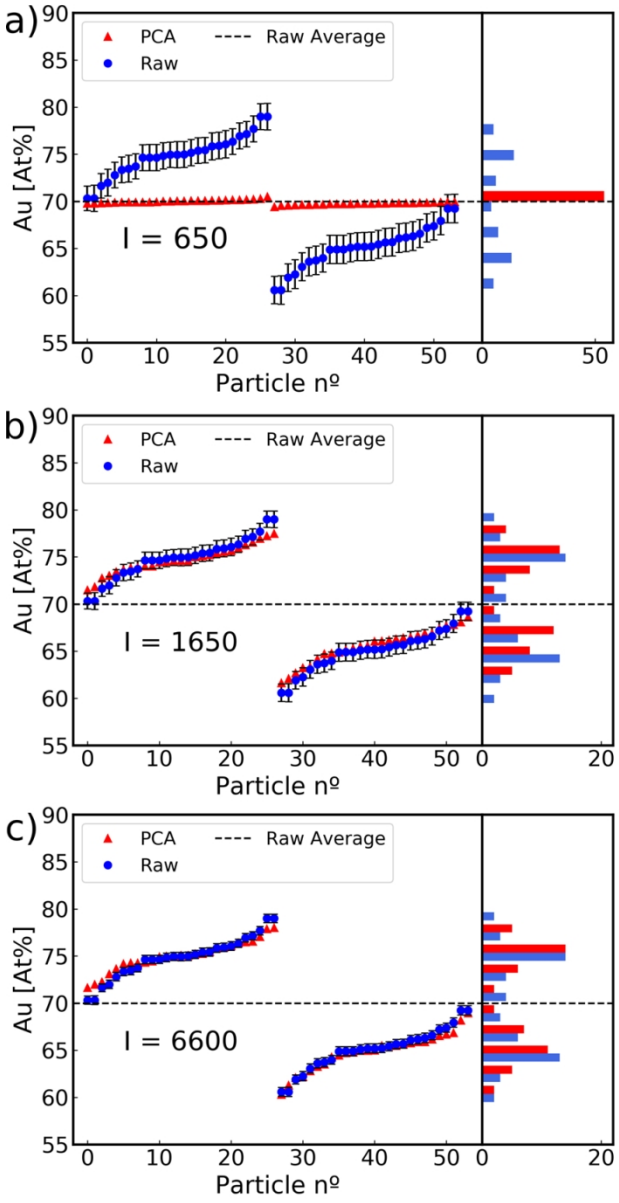

Figure 8. Effect of denoising on the assessment of chemical composition distribution deduced from simulated data sets including bimodal "true" distributions of chemical composition (population 1: C_(Au,1)^ =0.75 , population 2: C_(Au,2)^ =0.65, for both the distribution width is 〖v^*〗_(C_Au,STD)=0.02). The calculated chemical composition distribution is plotted at the right (see text for explanations).

317x604mm (72 x 72 DPI)

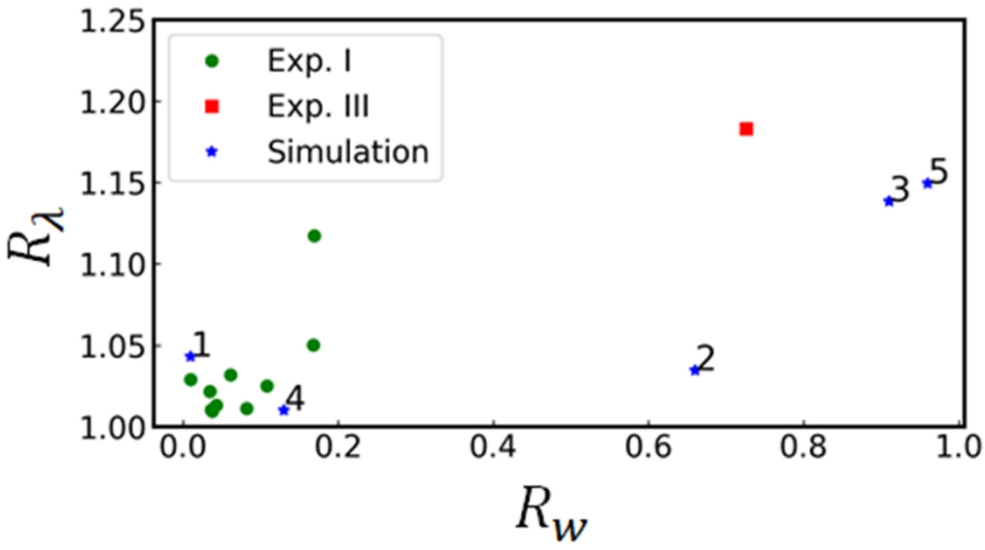

Figure 9. Plot showing the correlation of proposed bias estimators R_w and R_λ for Experiments I and III and simulations displayed in Fig. 6 (index 1-5 indicated simulation describes in Fig. 6a-6f respectively).

413x228mm (72 x 72 DPI)

Improving Quantitative EDS Chemical Analysis of alloy Nanoparticles by PCA Denoising: Part I, reducing reconstruction bias.

Murilo Moreira[1], Matthias Hillenkamp[1,2], Giorgio Divitini[3], Luiz H. G. Tizei[4], Caterina Ducati[3], Monica A. Cotta[1],Varlei Rodrigues[1,*], Daniel Ugarte[1,*]

1. *Instituto de Fisica "Gleb Wataghin", Universidade Estadual de Campinas-UNICAMP, 13083-859, Campinas - SP, Brazil*
2. *Institute of Light and Matter, Université Claude Bernard Lyon 1, CNRS, UMR5306, F-69622 Villeurbanne, France*
3. *Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB3 0FS, United Kingdom*
4. *Laboratoire de Physique des Solides, Université Paris-Saclay, CNRS, 91405 Orsay, France*

***Correspondence and requests for materials should be addressed to:**
D.U. (email: dmugarte@ifi.unicamp.br) or to V.R. (email: varlei@ifi.unicamp.br)

Figure S1.

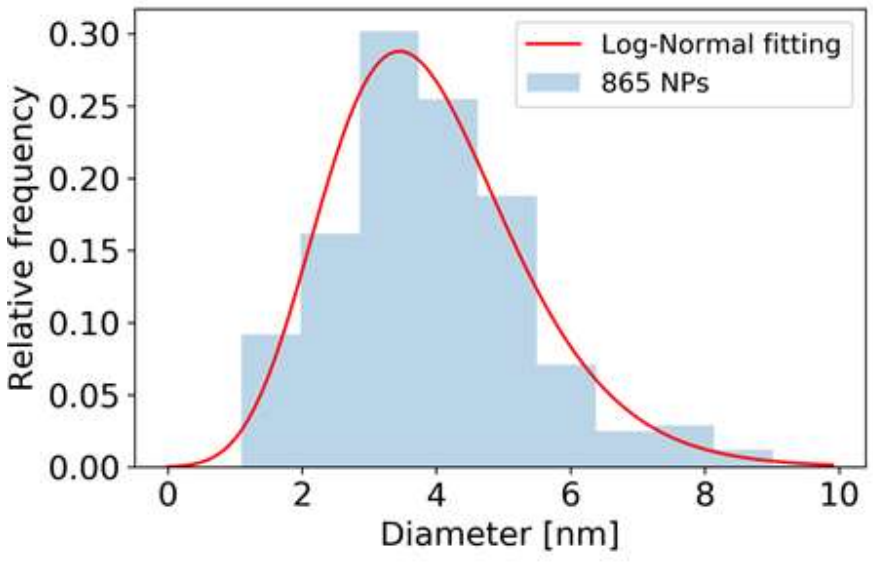

Figure S1. Typical size distribution of the alloy NP used in Experiment I, II and III as measured by TEM. The continuous curve represents a log-normal description of the distribution.

Figure S2.

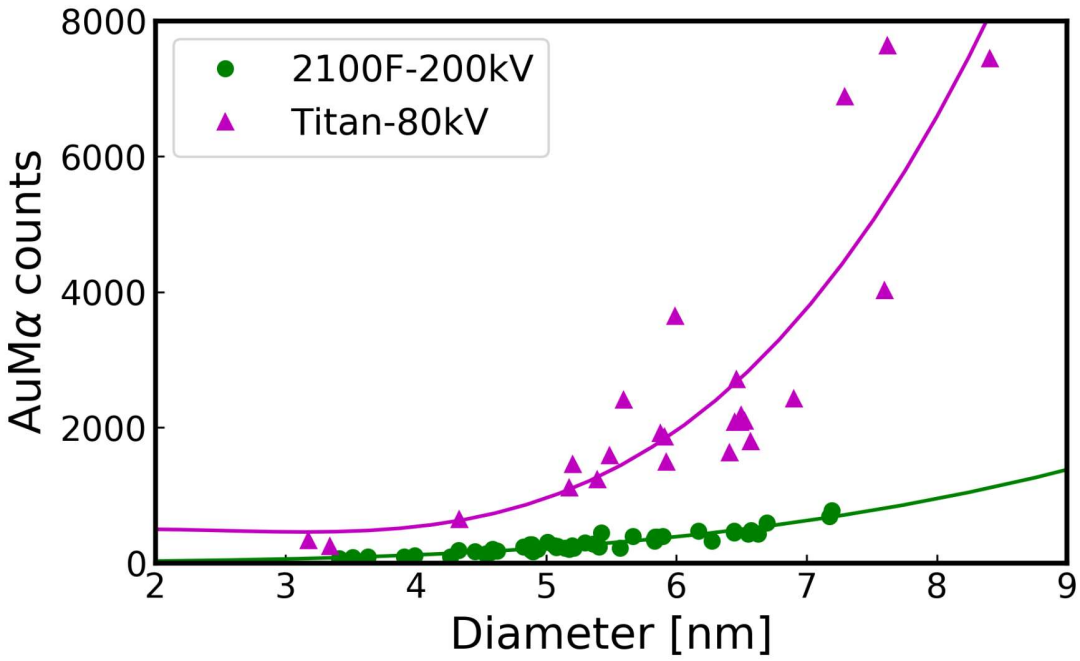

Figure S2. Integrated total counts per NP at the Au Lα peak (200 ms dwell time), as a function of NP diameter for different instruments. This allows a quick estimation of expected counts and Poisson noise for nanoparticle. A the mathematical expression describing the general tendency of the plots has been obtained by fitting the data. The observed counts as a function of NP diameter (D) can be estimated the equation below, derived from Figure S2.

for Titan-80kV, $I(D) = 38.3\ D^3 - 287.4\ D^2 + 670.7\ D$

for JEM2100F-200kV, $I(D) = 2.4\ D^3 - 6.9\ D^2 + 18.6\ D$

Figure S3.

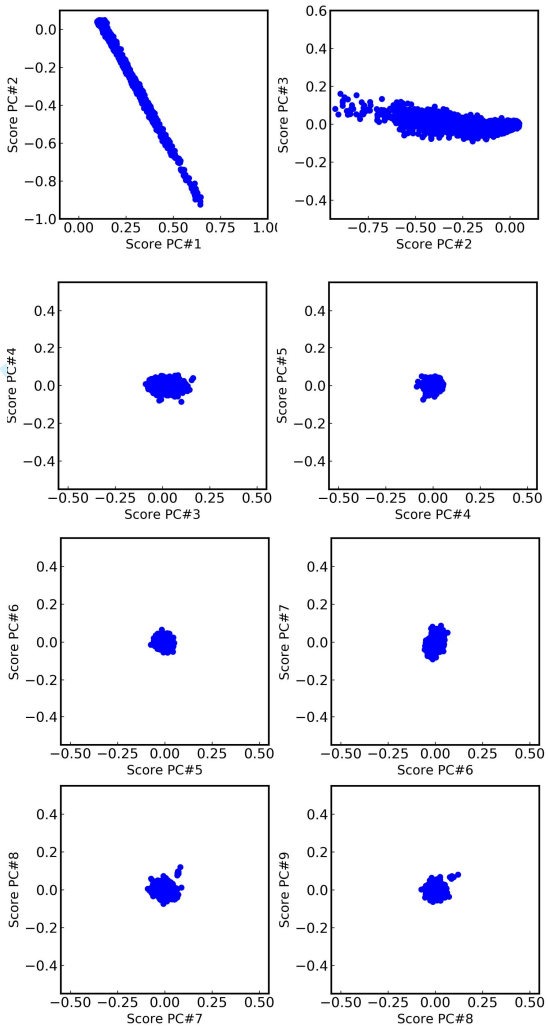

Figure S3. Cluster plot of scores for successive principal components obtained PCA processing a high SNR simulated EDS HSI of bimetallic NPs (Potapov, P., Lubk, A. (2019). Optimal principal component analysis of STEM XEDS spectrum images. *Adv. Struct. Chem. Imag.* ***5***, 4). Note that information-carrying component shows a structured anisotropic cloud (top), while noise components generate isotropic round clouds.